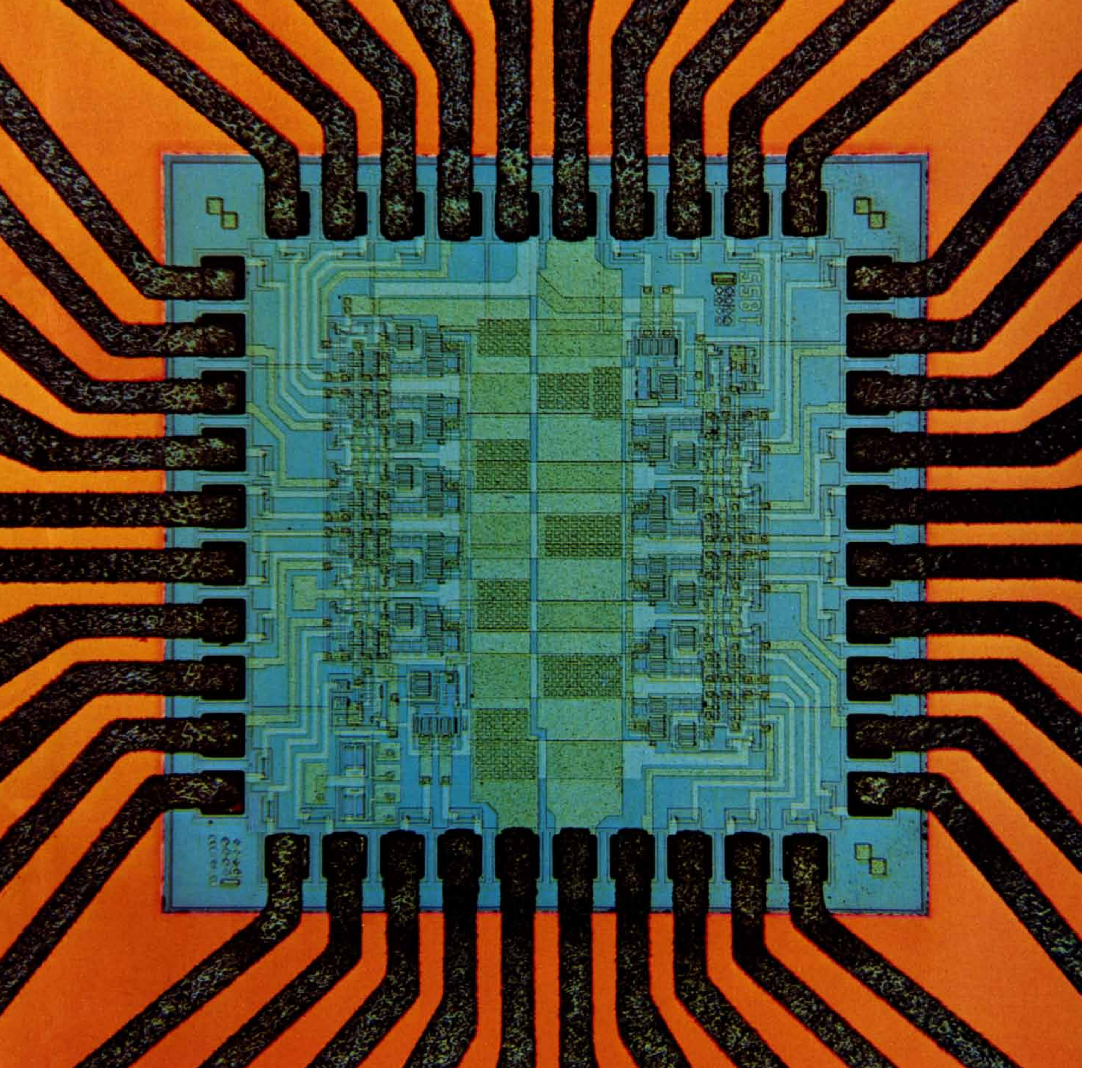
INVESTIGACION Y CIENCIA

Edición en español de SCIENTIFIC AMERICAN



MICROELECTRONICA

Noviembre 1977

200 Ptas.

Los espacios en gris
corresponden a publicidad
en la edición impresa

Número 14
Noviembre 1977

INVESTIGACION Y
CIENCIA
Edición en español de SCIENTIFIC AMERICAN

EDITA Prensa Científica, S.A.
Calabria, 235-239
Barcelona-15
ESPAÑA

Colaboradores de este número:

Asesoramiento y traducción:

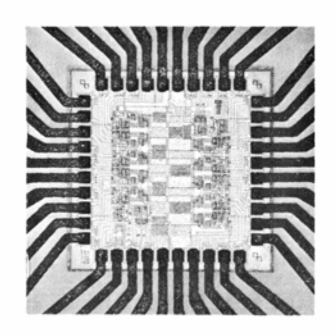
Alfredo Bautista: *Microelectrónica;* Santiago Rey: *Elementos del circuito microelectrónico;* Mauricio Wagensberg: *Integración a gran escala en los circuitos microelectrónicos;* Manuel Medina: *La fabricación de microcircuitos electrónicos;* Antonio Creus: *Memorias microelectrónicas;* Enrique Herrada: *Microprocesadores;* Depto. traducciones IBM España: *Microelectrónica y proceso de datos;* Antonio Creus: *Microelectrónica e instrumentación y control;* Juan Mulet: *Microelectrónica y comunicaciones;* Carme Basté: *Microelectrónica e informática;* Ana Ripoll: *Microelectrónica y ordenador de uso personal;* Luis Bou: *Juegos matemáticos;* José E. Myro y Alicia Menéndez: *Taller y laboratorio;* Gerardo Espinosa: *Libros.*

Ciencia y sociedad:

Bernardo Ancochea y Federico Mayor Zaragoza

Libros:

Philip Morrison



LA PORTADA

La fotografía de la portada simboliza el tema de este número monográfico de INVESTIGACION Y CIENCIA, dedicado a la microelectrónica, que es el arte de colocar complejos circuitos electrónicos en "pastillas" de silicio de unos seis milímetros de lado. La superficie azul, en el centro de la fotografía, es una pastilla única, un circuito lógico integrado de alta velocidad, fabricado por Texas Instruments para un ordenador de la firma Honeywell Information Systems. Rodean a la pastilla 40 conductores que la conectan con su entorno. Los conductores son de cobre recubierto con una capa de estaño. Se conectan por medio de un proceso automático, en el que se enlazan por acción del calor y la compresión a terminales de oro depositados en el borde de la pastilla. F. Goro tomó la fotografía en el laboratorio de Applied Microscopy de E. Leitz, Inc.

Suscripciones:

Prensa Científica, S.A.
Departamento de suscripciones
Calabria, 235-239
Barcelona-15 (España)

Distribución para España:

Distribuciones de Enlace, S.A.
Ausias March, 49
Barcelona-10 (España)

Distribución para los restantes países:

Editorial Labor, S.A.
Calabria, 235-239
Barcelona-15 (España)

Precio de venta de este número:

España: 200 pesetas
Extranjero: 2,9 U.S. $

Condiciones de suscripción:

España:
Un año (12 números): 1.540 pesetas
Extranjero:
Un año (12 números): 30 U.S. $

Ejemplar atrasado ordinario:
160 pesetas

Ejemplar atrasado extraordinario:
225 pesetas.

Solicitado control **OJD**

PROCEDENCIA DE LAS ILUSTRACIONES

Fotografía de la portada de Fritz Goro

Página	Fuente
6-7	Jon Brenneis
8	Fritz Goro *(izquierda)*, Fairchild Semiconductor *(derecha)*
9	Fritz Goro
10	Ben Rose
17	Fritz Goro
18-27	Gabor Kiss
31	Texas Instruments Incorporated
32-43	Gabor Kiss
47	Open Court Publishing
48	Editions du Seuil
54-55	Ben Rose
56	George V. Kelvin
57-59	Jon Brenneis
60	Jon Brenneis *(arriba)*, William Roth *(centro)*, Zilog Inc. *(abajo)*,
61	George V. Kelvin *(arriba)*, William Roth *(abajo)*,
62	George V. Kelvin
63	Mostek Corporation
64	George V. Kelvin
65-66	Jon Brenneis
69	Mostek Corporation
70-71	Allen Beechel
72	David A. Hodges *(arriba)*, Allen Beechel *(abajo)*,
73	Allen Beechel
74	Fairchild Semiconductor *(arriba)*, Texas Instruments Incorporated *(abajo)*
76-77	Allen Beechel
78	International Business Machines Corporation
79	Allen Beechel
81	Intel Corporation
82-84	Adolph E. Brotman
85	James Bell
88-90	Adolph E. Brotman
93-101	A. Alabau y A. Antoni Sellés
102	James Bell
106	Ralph Morse
107	Ralph Morse *(arriba)*
108	Ralph Morse
109	International Business Machines Corporation *(arriba)*,
110	Digital Equipment Corporation
115	James Bell
116	Allen Beechel
117	Ralph Morse
118	James Bell *(arriba)*, Allen Beechel *(abajo)*,
119	Ralph Morse
120	The Foxboro Company
121	Allen Beechel
125	James Bell
126	James Bell *(arriba)*, Fritz Goro *(abajo)*,
127	James Bell *(arriba)*, Fritz Goro *(abajo)*,
128-129	Ben Rose
130	James Bell *(arriba)*
131	Ben Rose
135-136	Fritz Goro
137	Alan D. Iselin
138	Ralph Morse *(arriba)*, TRW Inc. *(abajo)*,
139	James Bell *(arriba)*, Ben Rose *(abajo)*,
140	Digital Equipment Corporation *(arriba)*,
141-146	Alan D. Iselin
148	Evans & Sutherland Computer Corporation
150	James P. Mayer
151-160	Jon Brenneis
171-173	Michael Goodman

Dep. legal: B. 38.999-76
Fotocomposición: Fort, S.A.
Rosellón, 33 - Barcelona-15
Cayfosa. Santa Perpetua de Moguda. Barcelona.

Printed in Spain - Impreso en España

Los autores

ROBERT N. NOYCE ("Microelectrónica") preside Intel Corporation. Asistió al Grinnell College y al Instituto de Tecnología de Massachusetts, consiguiendo el doctorado en física electrónica por el M.I.T. en 1953. Trabajó en el equipo de investigación de Philco Corporation y luego en el Shockley Semiconductor Laboratory, en donde participó en el diseño y desarrollo de los transistores de silicio. En 1957 contribuyó a la fundación de la Fairchild Semiconductor Corporation, de la que durante 10 años fue director de investigación, vicepresidente y director general, sucesivamente. Dejó Fairchild en 1968 para cofundar Intel, compañía "dedicada a hacer realidad la integración a gran escala de los circuitos microelectrónicos".

JAMES D. MEINDL ("Elementos del circuito microelectrónico") es profesor de ingeniería eléctrica en la Universidad de Stanford y director de los laboratorios de electrónica de la misma. Graduado Ph. D. en ingeniería eléctrica por la Universidad de Carnegie-Mellon en 1958, trabajó durante dos años en la Westinghouse Electric Corporation, para pasar al Army Electronics Laboratory, en donde estuvo desde 1959 a 1967, ocupando diversos cargos. En 1967 ingresó, como profesor adjunto de ingeniería eléctrica en la Universidad de Stanford y actualmente trabaja en la investigación de la tecnología de los circuitos integrados en aplicación al desarrollo de equipos de ultrasonidos para diagnosis de tipo clínico.

WILLIAM C. HOLTON ("Integración a gran escala en los circuitos microelectrónicos") dirige el Advanced Components Laboratory de la compañía Texas Instruments Incorporated. Se graduó en física por la Universidad de Carolina del Norte, en 1952. Después de dos años y medio en la Marina, realizó trabajos de postgraduado en la Universidad de Illinois, obteniendo su doctorado en 1960. Ese mismo año se incorporó a la Texas Instruments como miembro del personal técnico. En 1967 fue nombrado director de la sección de electrónica cuántica, donde dirigió investigaciones y desarrollos en rayos láser y en óptica integrada. Accedió a su situación actual en 1973. Holton se responsabiliza de la dirección de la investigación y del trabajo de desarrollo de los circuitos integrados a gran escala, litografía por haz electrónico y por rayos X, dispositivos y representaciones de microondas.

WILLIAM G. OLDHAM ("La fabricación de microcircuitos eléctricos") es profesor de ingeniería eléctrica en la Universidad de California en Berkeley. Estudió en la Carnegie-Mellon University, recibiendo su Ph.D. en 1963. Después de un año en el equipo de investigación de Siemens Corporation, en Alemania, se incorporó al centro de Berkeley. En 1969, durante el año sabático, estuvo en la Technische Universität de Munich, y en 1974 se trasladó a Intel Corporation, donde dirigió el desarrollo de la memoria microelectrónica para computador con una capacidad de almacenamiento de 16.000 bits. El campo de investigación actual de Oldham se centra en las técnicas de la fabricación de dispositivos microelectrónicos de dimensiones micrométricas.

DAVID A. HODGES ("Memorias microelectrónicas") es profesor de ingeniería eléctrica y ordenadores en la Universidad de California en Berkeley. Se licenció en ingeniería eléctrica por la Universidad de Cornell y se doctoró en ingeniería eléctrica por la de Berkeley, en 1966. A partir de entonces trabajó en los Bell Laboratories, primero en el área de componentes en Murray Hill, N. J. y a continuación como jefe del departamento de investigación de elementos de sistemas en Holmdel, N.J. En el año 1970, entró a formar parte de la facultad de Berkeley y, a partir de entonces, trabajó como consultor de industrias de semiconductores y de equipamiento de comunicaciones.

HOO-MIN D. TOONG ("Microprocesadores") es profesor adjunto de ciencias de la computación e ingeniería en el Instituto de Tecnología de Massachusetts (M.I.T.) y también profesor adjunto de ciencias empresariales en la Sloan School of Management del M.I.T. Hizo sus estudios en el M.I.T., donde recibió los títulos de ingeniero eléctrico en 1969 y el Ph.D. en 1974. A continuación formó parte del claustro de profesores del M.I.T. Además de impartir cursos de computadores para estudiantes de ingeniería, ha dado conferencias sobre microprocesadores por Estados Unidos, Europa y el Extremo Oriente. Toong tiene a su cargo el laboratorio de sistemas digitales del M.I.T., donde sus investigaciones se han centrado sobre problemas de soporte lógico y de soporte físico en sistemas de multiprocesador y en sistemas distribuidos. También se dedica a aplicar la tecnología del microprocesador a métodos para la conservación de la energía y el control de su utilización.

ANTONIO ALABAU y JOAN FIGUERAS ("Microcomputadores") son, respectivamente, catedrático de ordenadores de la Escuela Técnica Superior de Ingenieros de Telecomunicación y profesor agregado de electrónica de la Escuela Técnica Superior de Ingenieros Industriales, de la Universidad Politécnica de Barcelona. Alabau nació en Alfafar (Valencia), cursó estudios de ingeniería de telecomunicación en Madrid, realizó trabajos de investigación en el Laboratoire d'Automatique et d'Analyse des Systèmes del C.N.R.S. en Toulouse, doctorándose en la Universidad Paul Sabatier de dicha ciudad. Figueras, que nació en Valls (Tarragona), cursó los estudios de ingeniería industrial en Barcelona y realizó trabajos de investigación en el Systems Laboratory de la Universidad de Michigan; en 1971 fue profesor visitante en la Universidad de Chile, en Santiago; se doctoró en el departamento de electricidad y electrónica de la Universidad de Michigan (Ann Arbor). Desde 1972, ambos realizan actividades de investigación, desarrollo y docencia, respectivamente, en los Laboratorios de Ordenadores (E.T.S.I.T.) y en el Laboratorio de Investigación y Desarrollo de Electrónica (E.T.S.I.I.B.), pertenecientes ambos centros a la Universidad Politécnica de Barcelona.

LEWIS M. TERMAN ("Microelectrónica y proceso de datos") es miembro del estamento directivo de investigación del Thomas J. Watson Research Center, de la International Business Machines Corporation. Estudió en la Universidad de Stanford y obtuvo el grado de doctor en 1961. Después ingresó en el Watson Research Center, en donde trabajó en diseño de lógica de ordenador, y en 1963 fue nombrado director de un grupo científico dedicado a la investigación de la tec-

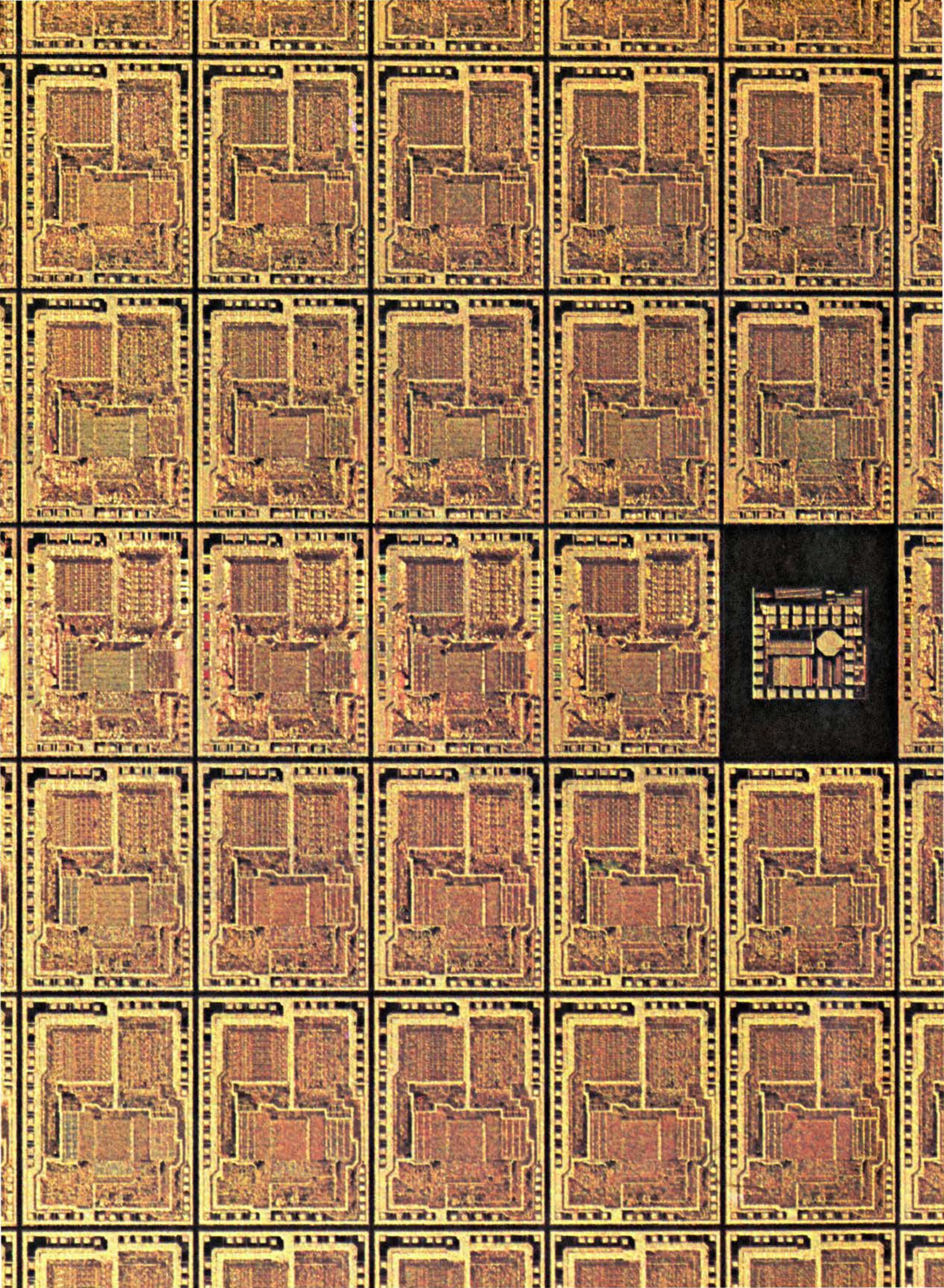
nología de las memorias magnéticas. Dos años más tarde, se dedicó al estudio del uso de dispositivos microelectrónicos para las memorias de los ordenadores, convirtiéndose en uno de sus primeros defensores. Normalmente se dedica a la investigación de dispositivos incorporados a las memorias y al estudio de los efectos de la integración a gran escala en el diseño de sistemas de ordenador.

BERNARD M. OLIVER ("Microelectrónica e instrumentación y control") es vicepresidente de investigación y desarrollo de la Hewlett-Packard Company y miembro de su consejo de administración. Se educó en la Universidad de Stanford y en el Instituto de Tecnología de California, en donde se licenció en 1936. Tras un año de estudio en Alemania, regresó al Instituto de Tecnología de California y obtuvo su doctorado, magna cum laude, en 1940. Se trasladó luego a los Bell Laboratories; en los siguientes 12 años participó en el desarrollo del radar de seguimiento automático, en la transmisión por televisión, en la teoría de la información y en los sistemas de codificación. En 1952 entró en Hewlett-Packard como director de investigación; pasó a su cargo actual en 1957. Oliver posee más de 50 patentes en el campo de la electrónica, y ha actuado como miembro de la Comisión Presidencial para el Sistema de Patentes. En 1966 fue elegido para la National Academy of Engineering y en 1973 para la National Academy of Sciences.

JOHN S. MAYO ("Microelectrónica y comunicaciones") es vicepresidente de tecnología electrónica en los Bell Laboratories. Nacido en Carolina del Norte, se educó en su universidad estatal, alcanzando el doctorado en ingeniería eléctrica en 1955. Entró entonces en los Bell Laboratories para realizar inicialmente investigaciones sobre ordenadores digitales transistorizados y sobre el uso de los ordenadores en sistemas de defensa. También desarrolló el método de transmisión de información en forma digital, conocido por modulación por impulsos codificados. Más tarde, aplicó la técnica al diseño del decodificador de mando y la unidad de conmutación del satélite de comunicaciones Telstar y al desarrollo del sistema "Picture-phone". Mayo pasó a ocupar su actual puesto en 1975. Es responsable de la dirección del desarrollo de componentes electrónicos, fuentes de energía, fuentes de alimentación y de los diseños de acondicionamiento de edificios para las campañías de explotación telefónica del Bell System.

IVAN E. SUTHERLAND y CARVER A. MEAD ("Microelectrónica e informática") son, respectivamente, director y profesor en el Departamento de Informática del Instituto de Tecnología de California. Sutherland se formó en el Instituto Carnegie de Tecnología, en el Instituto de Tecnología de California y en el de Massachusetts. Recibió el doctorado en ingeniería eléctrica en 1963, en el M.I.T. Entró en el Army Signal Corps, para ser destinado a la National Security Agency, donde diseñó una nueva línea de equipo de visualización de calculadores. En 1964, Sutherland fue nombrado director de técnica de proceso de la información en el ARPA ("Advanced Research Projects Agency") del Departamento de Defensa, donde permaneció hasta su retiro del Ejército. Al dejar el ARPA en 1966, pasó a ser profesor adjunto de ingeniería eléctrica en la Universidad de Harvard. Dos años después, se trasladó a Salt Lake City para presidir la Evans y Sutherland Computer Corporation, de la que fue cofundador. Siguió investigando en la Universidad de Utah, en donde colaboró en el diseño de un sistema gráfico por ordenador capaz de dar imágenes en color con perspectiva y sombreado realistas. Se incorporó al Instituto de Tecnología de California en 1976. Mead se formó en el Instituto de Tecnología de California, en donde se doctoró en ingeniería eléctrica en 1959. Empezó a investigar en la física de los dispositivos de estado sólido con una breve incursión en el campo de la biofísica de las membranas nerviosas. Empezó a interesarse por los problemas de gestión de los sistemas con integración a gran escala, al estudiar las limitaciones físicas de un transistor operativo de tamaño mínimo.

ALAN C. KAY ("Microelectrónica y ordenador de uso personal") es director científico y jefe del Learning Research Group perteneciente al Xerox Palo Alto Research Center. Se graduó en matemáticas por la Universidad de Colorado en Boulder, y después de una corta carrera como guitarrista profesional de jazz, estudió informática en la Universidad de Utah, obteniendo el título de Doctor en 1969. Ha sido investigador asociado y conferenciante en el proyecto de inteligencia artificial de la Universidad de Stanford. Se trasladó a Xerox Palo Alto Research Center en 1971.



Microelectrónica

Presentación de un número monográfico consagrado a la revolución microelectrónica, merced a la cual la reunión de muchos elementos en "pastillas" ha aumentado las posibilidades de los dispositivos electrónicos

Robert N. Noyce

La evolución de la tecnología electrónica en la última década se ha desplegado con tanta rapidez que en más de una ocasión se ha hablado de ella como de una revolución. ¿Merece ese nombre? Opino que sí.

Cierto es que lo que hemos visto ha sido de algún modo una evolución cuantitativa continuada: componentes electrónicos cada vez más pequeños realizando funciones electrónicas de complejidad creciente a velocidades siempre más altas y costes cada vez menores. Pero también ha habido una verdadera revolución, a saber: un cambio cualitativo en la tecnología, el circuito microelectrónico integrado, que ha provocado, a su vez, un cambio cualitativo en las posibilidades humanas.

No es ninguna exageración afirmar que la mayoría de los descubrimientos tecnológicos de la última década han dependido de la microelectrónica. Dispositivos de control y sensores, pequeños y seguros, son los elementos esenciales que han llevado hombres a la luna y han explorado Marte, por no hablar del papel similar representado en los medios de defensa intercontinentales que dominan la política mundial. Los dispositivos microelectrónicos son también la base de nuevos productos, desde los satélites de comunicación hasta las calculadoras de bolsillo y los relojes digitales. Algo más sutil, pero quizá más significativo, es la influencia de la microelectrónica en el ordenador. La capacidad del ordenador para almacenar, procesar y presentar la información ha experimentado un gran desarrollo. Además, para muchos cometidos, el ordenador se está trasladando a los lugares que demandan su uso o donde se aplican sus resultados: junto al operador, el instrumento o el dispositivo de control industrial.

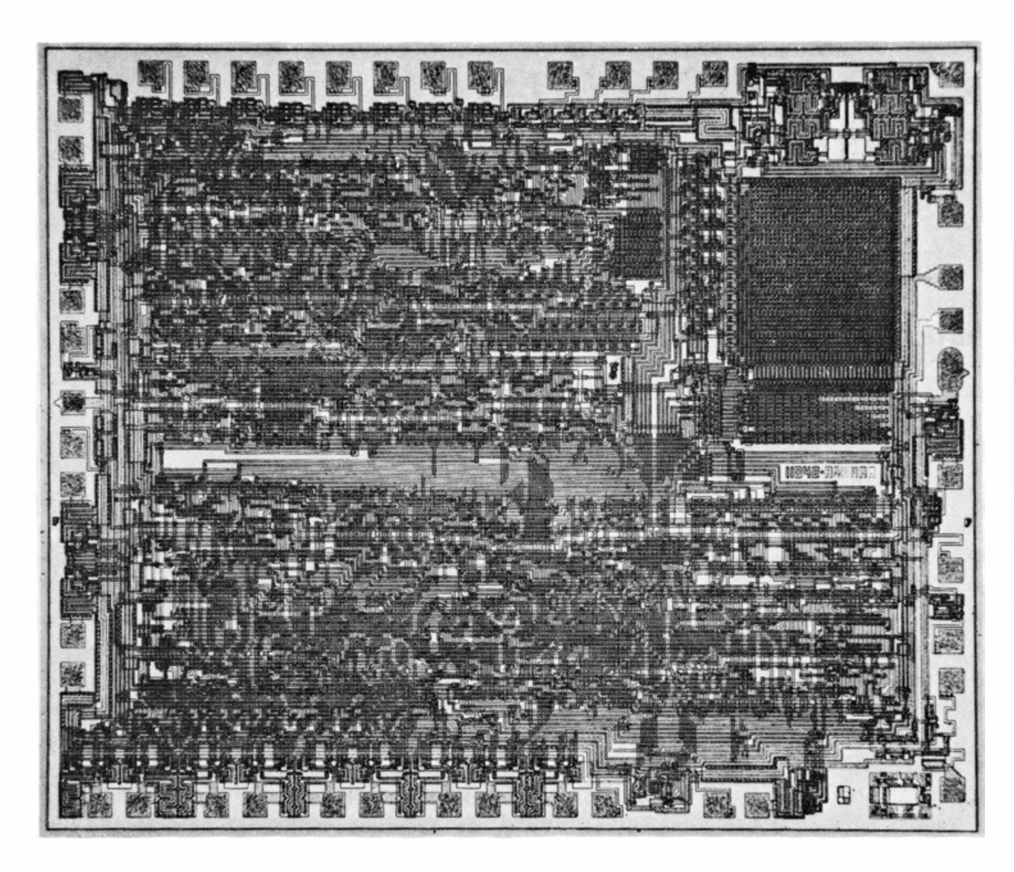
La revolución microelectrónica se halla lejos de haber cubierto su trayectoria. Estamos aprendiendo aún cómo aprovechar el potencial del circuito integrado mediante el desarrollo de nuevos circuitos, cuyo rendimiento puede mejorarse todavía en otro orden de magnitud. Y apenas si nos hemos adentrado en las implicaciones intelectuales y sociales del ordenador de uso personal, que puede proporcionar el acceso, por parte del individuo, a grandes almacenamientos de información y la posibilidad de aprender, incrementar y comunicar con otros dicha información.

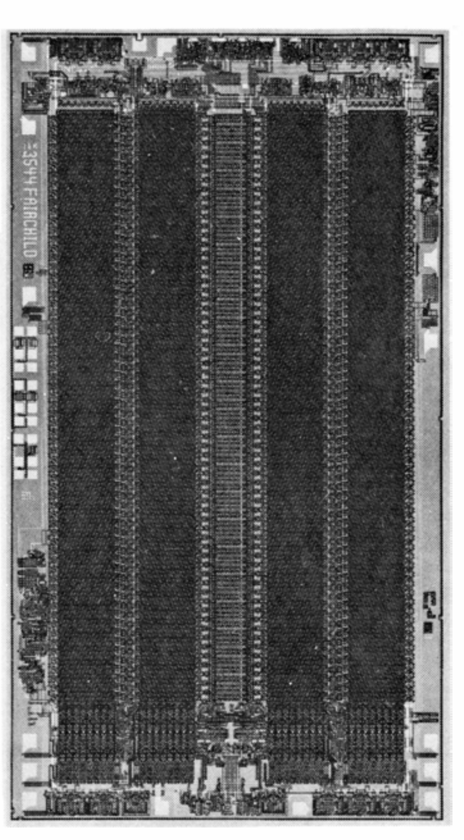
Este número de *Investigación y Ciencia,* dedicado a la microelectrónica, examina la naturaleza de los elementos que componen los circuitos microelectrónicos, el diseño y fabricación de los circuitos integrados a gran escala, una amplia extensión de sus aplicaciones y algunas de sus implicaciones para el futuro. Aquí, mi intención primera es mostrar cómo la evolución de la microelectrónica ilustra la interacción constante entre la tecnología y la economía. El exiguo tamaño de los dispositivos microelectrónicos ha resultado de interés en muchas aplicaciones, pero la principal incidencia de esta nueva tecnología ha estribado en lograr que las funciones electrónicas puedan reproducirse más, sean más seguras y más baratas. A cada progreso técnico ha correspondido un nuevo abaratamiento, y el descenso de los costes ha servido de incentivo a un amplio campo de aplicaciones: la búsqueda de avances tecnológicos ha venido motivada por la competencia económica y se ha visto compensada por la remuneración económica.



LOS DISPOSITIVOS MICROELECTRONICOS que están agrupados en la página opuesta son microprocesadores, cada uno de los cuales equivale a la unidad central de procesos de un pequeño ordenador. La fotografía muestra una porción de "oblea", que es una fina capa de silicio sobre la cual se han fabricado los dispositivos, aumentada 10 diámetros. Sobre este epígrafe se muestra un dispositivo individual, o pastilla, a tamaño real (0,417 por 0,564 cm). Después que los dispositivos se han fabricado sobre la oblea y se han comprobado, se separan, se les dota de conexiones externas y se empaquetan. Este microprocesador es el 8085 de la Intel Corporation, dispositivo de objetivos generales que tiene 6200 transistores y está capacitado para ejecutar 770.000 instrucciones por segundo. La pastilla más pequeña de la derecha se utiliza para someter a prueba la oblea en las diferentes etapas de su elaboración.

Todo comenzó con el desarrollo, hace 30 años, del transistor, es decir, de un amplificador pequeño y de baja potencia que reemplazó a la válvula de vacío, de mayor consumo y tamaño. El advenimiento, casi simultáneamente, del ordenador digital de programa almacenado le abrió al transistor un gran mercado potencial. La conjunción entre un nuevo componente y una nueva aplicación generó un crecimiento explosivo de ambos. El ordenador fue el mercado ideal para el transistor y para los circuitos integrados de estado sólido, que el transistor trajo como consecuencia, mercado mucho más fecundo que el que podía haber encontrado en las aplicaciones tradicionales de la electrónica al campo de las comunicaciones. Ello es así porque los sistemas digitales precisan gran número de circuitos activos, si los comparamos con los sistemas dotados de amplificación analógica, como la radio. En la electrónica digital, un elemento determinado se halla en uno de estos dos estados: activado o inactivado; lo cual depende de la entrada. Por muchos elementos que se conecten, la salida de los mismos continuará estando sencillamente activada o inactivada. La ganancia de la etapa individual es la unidad, de modo que, aunque se encadenen en cascada varias etapas, la ganancia seguirá siendo la unidad. Por otra parte, los circuitos





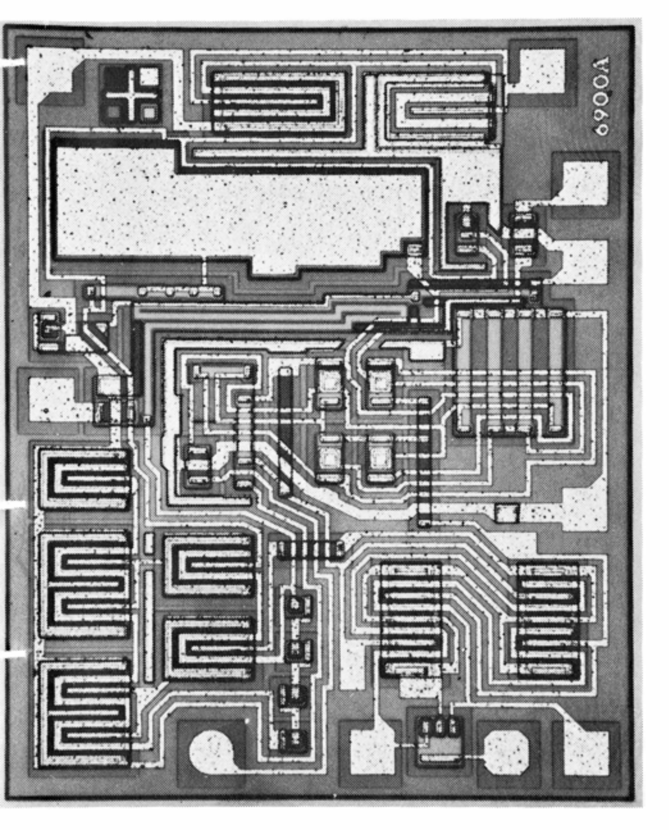
CUATRO PASTILLAS DISTINTAS, que sugieren la amplia gama de objetivos cubiertos por los dispositivos microelectrónicos y las características típicas de esos objetivos. A la izquierda, se halla un circuito integrado a gran escala: un microprocesador fabricado por la Rockwell International Corporation. Además de los elementos lógicos de una unidad central de procesos, incorpora una circuitería de entrada-salida y una memoria de sólo lectura *(área rectangular, arriba a la derecha).* **El segundo de la izquierda es otro circuito integrado a gran escala: una memoria de semiconductores fabricada por Fairchild Semiconductor. Su apariencia regular es característica de los**

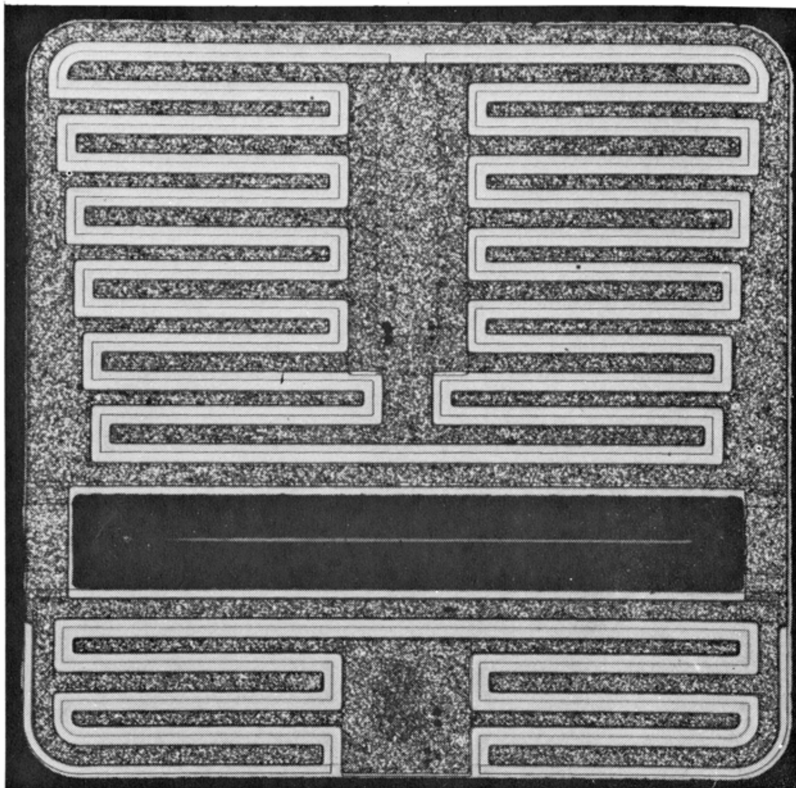
analógicos requieren amplificación de la entrada. Como la ganancia de cada amplificador puede valer típicamente 10, sólo pueden conectarse en serie o cascada unas cuantas etapas antes de que se alcance el límite práctico de los niveles de tensión para elementos microelectrónicos. El sistema analógico no puede manejar, por tanto, gran cantidad de microcircuitos, mientras que el sistema digital sí los precisa; una calculadora de bolsillo contiene 100 veces más transistores que una radio o un receptor de televisión.

A pesar de la compatibilidad esencial entre microelectrónica y ordenador, el hecho histórico es que los primeros esfuerzos encaminados a miniaturizar los componentes electrónicos no se debieron a los expertos en ingeniería de ordenadores. El tremendo potencial del ordenador digital no se apreció de inmediato; los mismos creadores del primer ordenador creyeron que con unos cuatro ordenadores quedarían cumplidas las necesidades mundiales en lo referente al cálculo. Sin embargo, distintos proyectos de misiles y satélites pidieron instalar en sus equipos unos sistemas electrónicos complejos, en los cuales estuvieran muy restringidas las exigencias de tamaño, peso y potencia. Fue así como nacieron los esfuerzos de miniaturización: instados por oficinas militares y espaciales.

Al principio se pretendió miniaturizar los componentes convencionales. Entre los programas iniciales se contaba el "Proyecto Tinkertoy", del National Bureau of Standards, cuyo objetivo era encapsular distintos componentes electrónicos en un modelo normalizado: una forma rectangular que pudiera empaquetarse de manera compacta, en vez de la forma cilíndrica tradicional. En otra perspectiva se encontraba "la ingeniería molecular". El ejemplo del transistor como sustituto de la válvula de vacío sugirió que podrían idearse sustitutos similares: podrían descubrirse o desarrollarse nuevos materiales que, por su naturaleza de estado sólido, permitieran efectuar funciones electrónicas, distintas de la amplificación, dentro de un sólido monolítico. Estos intentos fracasaron, pero divulgaron la demanda de miniaturización y las recompensas potenciales que comportaría el éxito en el desarrollo de alguna forma de microelectrónica. Gran parte de la comunidad técnica se halló comprometida en la búsqueda de una solución al problema, ya que estaba claro que al afortunado inventor le esperaba un mercado dispuesto.

El circuito integrado de semiconductores proporcionó finalmente la solución. Se trataba de una serie de ideas que habían empezado a tomar forma a los pocos años de la invención del transistor. Varios investigadores observaron que se podían aprovechar las características de ciertos semiconductores, como el germanio y el silicio, que habían sido utilizados para fabricar el transistor. La resistencia propia del semiconductor y la capacidad de las uniones entre las regiones positiva *(p)* y negativa *(n)* que podían crearse en él, podían combinarse con transistores dentro del mismo material para realizar un circuito completo de resistencias, condensadores y amplificadores (véase "Elementos del circuito microelectrónico", de James D. Meindl, en este número). En 1953 Harwick Johnson, de la Radio Corporation of America, patentó un oscilador por desplazamiento de fase incorporado en una





dispositivos de memoria. El tercero de la izquierda es un circuito integrado de pequeña escala: un "amp. op.", amplificador operacional, fabricado por RCA Corporation. Se trata de un amplificador variable, cuya salida está ligada a su entrada, y así sirve como bloque adaptable en distintas aplicaciones de realimentación. En la ilustración situada a la derecha, en esta página, se aprecia un dispositivo de potencia: es un "transistor Darlington" de alta ganancia, fabricado por Motorola Semiconductor Products Inc. En dicho dispositivo, dos transistores proporcionan suficiente potencia para activar la circuitería de deflexión horizontal propia de un receptor de televisión.

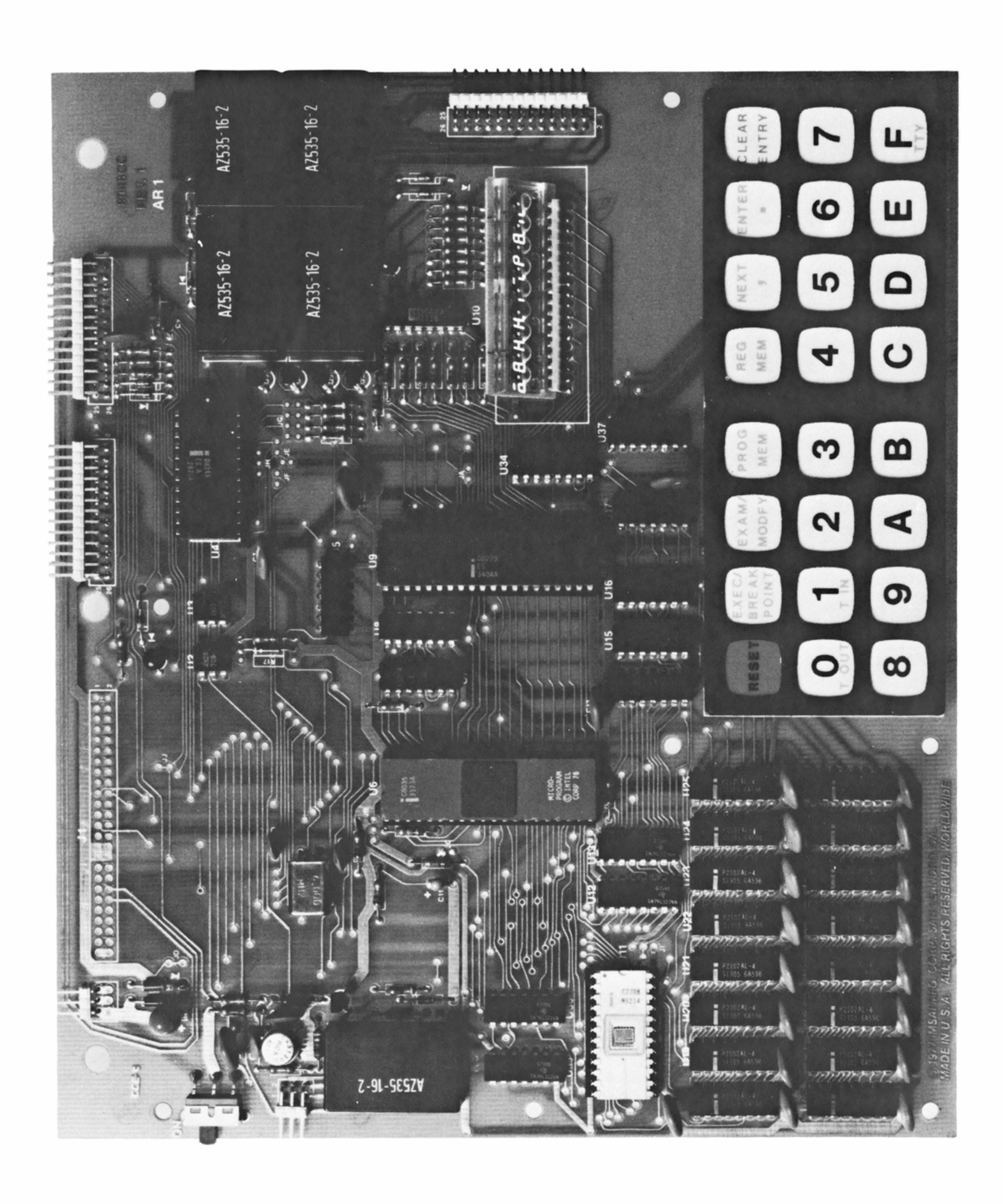
pieza de germanio mediante la técnica mencionada. G. W. A. Dummer, del Royal Radar Establishment (Inglaterra), Jack S. Kilby, de Texas Instruments Incorporated, y Jay W. Lathrop, de Diamond Ordnance Fuze Laboratories, generalizaron la idea.

Se precisaron, no obstante, varios avances clave antes de que pudiera comprobarse el apasionante potencial de los circuitos integrados. A mediados de los años 50, los ingenieros aprendieron a definir la configuración superficial de los transistores por medio de la fotolitografía, y desarrollaron el método de la difusión del estado sólido para inyectar las impurezas que crean las regiones *p* y *n*. El procesamiento en bloque de muchos transistores sobre una delgada "oblea", seccionada a partir de un cristal grande de germanio o silicio, comenzó a desplazar la técnica pionera del procesamiento individual de los transistores. Los cientos de miles de transistores registrados con precisión, que podían fabricarse sobre una sola oblea, tenían aún que ser separados físicamente, ensamblados individualmente con cables finos en el interior de una cápsula protectora y posteriormente incorporados en circuitos electrónicos.

El circuito integrado, tal como lo concebimos y desarrollamos en Fairchild Semiconductor en 1959, consigue la separación e interconexión de los transistores y demás elementos del circuito, eléctrica en vez de físicamente. La separación se obtiene introduciendo diodos *pn*, o rectificadores, que permiten que la corriente fluya en una sola dirección. Kurt Lehovec, de la Sprague Electric Company, patentó la técnica. Los elementos del circuito se interconectan mediante una película conductora de metal evaporado, que se fotograba para conseguir la apropiada red de conexiones. Se precisa una capa aislante para separar el semiconductor subyacente de la película de metal, excepto allí donde se desee que estén en contacto. El proceso que efectúa este aislamiento fue desarrollado por Jean Hoerni en Fairchild en 1958, año en que él inventó el transistor planar: sobre la superficie de la oblea se forma una delgada capa de dióxido de silicio, uno de los mejores aislantes conocidos, tras haber sido tratada y antes de que el metal conductor se evapore.

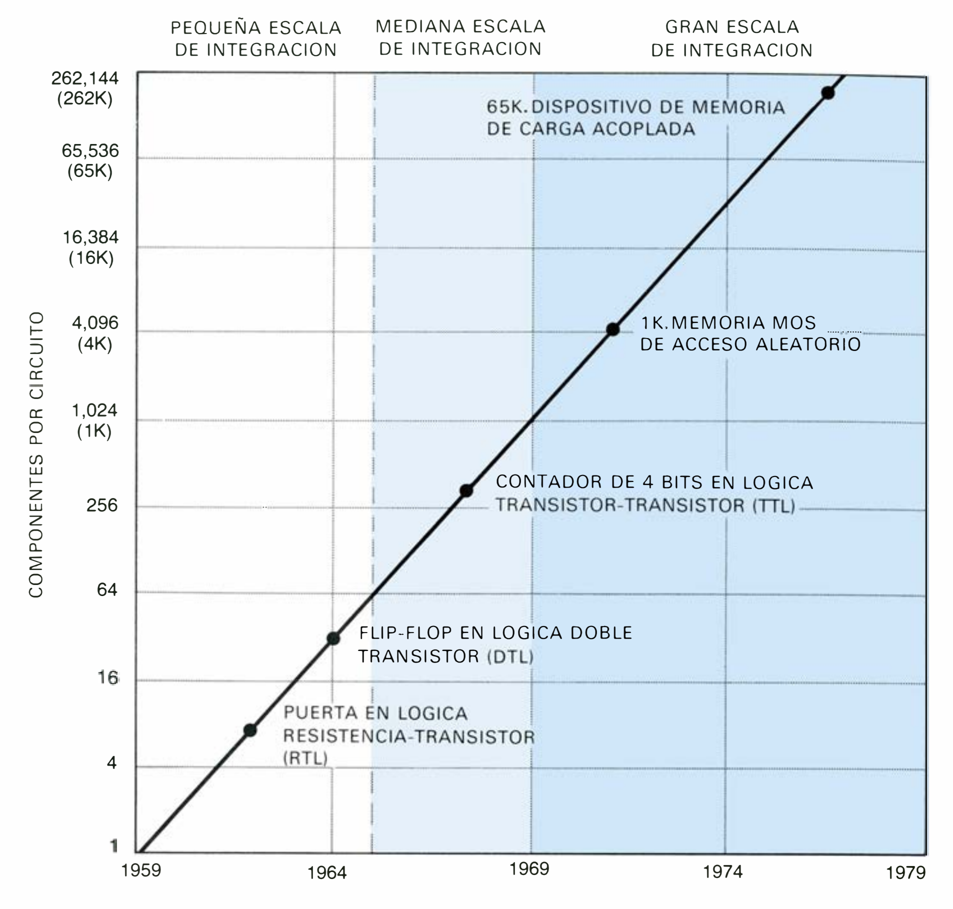
A partir de entonces se han ideado técnicas adicionales que brindan mayor flexibilidad al diseñador de circuitos, pero los métodos básicos estaban disponibles ya en 1960; se había inaugurado la era del circuito integrado. El progreso posterior ha resultado sorprendente, incluso para aquellos de nosotros que nos hallábamos comprometidos íntimamente con el despliegue tecnológico. Un solo circuito integrado en una pastilla de unos 0,6 centímetros puede englobar ahora más elementos electrónicos que la pieza más compleja de equipo electrónico que pudiera construirse en 1950. El microordenador actual, a un precio de unos 300 dólares, tiene más capacidad de cálculo que el primer gran ordenador electrónico, el ENIAC. Veinte veces más rápido, dotado de una memoria mayor, miles de veces más seguro, consume la potencia de una bombilla (y no la de una locomotora) y ocupa 1/30.000 del volumen, costando diez mil veces menos. Puede adquirirse por correo o en cualquier tienda del ramo.

En 1964, advirtiendo que, desde la producción del transistor planar en 1959, el número de elementos de los circuitos



EL MICROORDENADOR "IMSAI" 8048, fabricado por IMSAI Manufacturing Corporation, reúne circuitos integrados junto a otro tipo de dispositivos sobre una tarjeta de 20 por 25 centímetros. El microprocesador es la pastilla cuadrada de color gris claro que se encuentra en el centro, a un tercio de camino empezando desde la base. Las 16 pastillas *(abajo, a la derecha)* **constituyen una memoria de acceso aleatorio de 2K octetos (16384 bits o dígitos binarios). El encapsulado claro, próximo a ellas, proporciona 2K octetos programables de memoria de sólo lectura, además de la memoria de programa de 1K octeto que lleva el microprocesador. Nueve diodos emisores de luz** *(a la izquierda del teclado)* **proporcionan un monitor visual alfanumérico. Es éste un ordenador de control que puede, por ejemplo, gobernar una operación y accionar los interruptores controlándolos por medio de relés de alta corriente** *(las grandes cajas negras numeradas).*

integrados avanzados se había ido duplicando anualmente, Gordon E. Moore, que era por entonces director de investigación en Fairchild, fue el primero en prever el progreso futuro del circuito integrado. Sugirió que cada año continuaría duplicándose su complejidad. Hoy, con circuitos que contienen 2^{18} (262.144) elementos disponibles, no hemos detectado todavía ninguna desviación significativa de la ley de Moore. Ni hay señal de que el proceso se vaya a hacer más lento, aunque, en definitiva, resulta inevitable una desviación del crecimiento exponencial. La tecnología se halla lejos de los límites fundamentales que le imponen las leyes de la física: una mayor miniaturización parece estar menos condicionada por las leyes físicas que por las económicas.



EL NUMERO DE COMPONENTES en los circuitos integrados más avanzados se ha doblado anualmente desde 1959, año en que se desarrolló el transistor planar. El profesor Gordon E. Moore, del Fairchild Semiconductor, señaló la tendencia en 1964 y predijo su continuidad.

El crecimiento de la industria microelectrónica ilustra hasta qué punto las inversiones en el campo de la investigación pueden resultar en la creación de oportunidades empresariales, puestos de trabajo y mercados de exportación. Después de la introducción del circuito integrado a principios de los años 60, el consumo mundial total de circuitos integrados creció rápidamente, alcanzando un valor de casi mil millones de dólares en 1970. En 1976 el consumo mundial superó la triplicación de la cifra con los 3500 millones de dólares recabados. De este total, las compañías norteamericanas produjeron más de 2500 millones, alrededor del 70 por ciento, del cual mil millones se exportaron a clientes extranjeros. Pero el impacto registrado en la industria electrónica trasciende el marco escueto de esos números. En los equipos electrónicos, menos del 10 por ciento del valor se debe a los propios circuitos integrados: un miniordenador de 10.000 dólares no asigna ni siquiera 1000 de ellos a los circuitos electrónicos integrados, y un aparato de televisión que valga 300 dólares no cede ni siquiera 30. La mayoría de las industrias electrónicas del mundo (que representan un total de 80.000 millones) dependen, de alguna manera, de los circuitos integrados.

La sustitución de componentes discretos por dispositivos microelectrónicos reduce los costos; y ello, no sólo debido a que los dispositivos en sí sean más baratos, sino por muchas otras razones. Primero, el circuito integrado contiene muchas de las interconexiones que anteriormente eran necesarias, y, por tanto, ahorra trabajo y materiales. Las interconexiones de los circuitos integrados son mucho más fiables que las soldaduras y los conectores, lo cual se traduce en ahorro de mantenimiento. Como los circuitos integrados son mucho más pequeños y consumen bastante menos potencia que los componentes que han desplazado, permiten ahorrar en estructura de soporte (cabinas y bastidores), en transformadores de potencia y en dispositivos de ventilación. Durante la producción se necesitan menos comprobaciones intermedias, porque el comportamiento correcto de los complejos circuitos integrados está ya asegurado. Finalmente, el usuario economiza espacio, potencia de operación y refrigeración para el equipo. Todo esto es un modo de decir que, aun cuando los circuitos integrados sólo fueran equivalentes en coste a los componentes que han desplazado, otras ventajas motivarían el uso de circuitos integrados más elaborados y en menor cantidad conforme estuvieran disponibles.

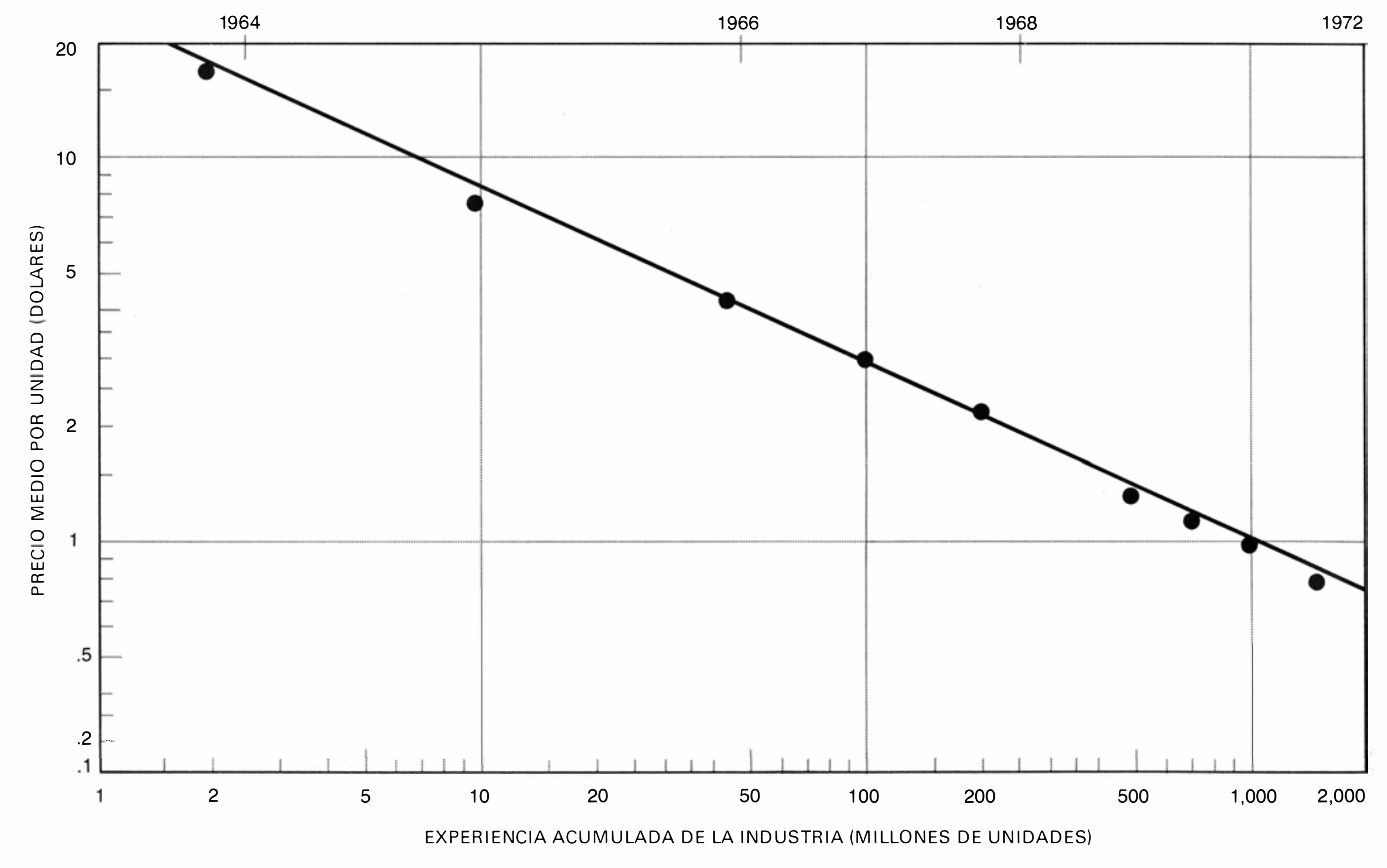
La característica más notable de la industria microelectrónica ha sido el descenso rápido y persistente en el coste de una función electrónica dada. La calculadora de bolsillo proporciona un claro ejemplo. Su coste ha disminuido en un factor de 100 durante la última década. Parte del rápido descenso de costes puede contabilizarse en función de la "curva de aprendizaje": cuanto más experiencia tiene una industria, tanto más eficaz llega a ser. La mayoría de las industrias reducen sus costes (en dólares constantes) entre un 20 y un 30 por ciento cada vez que su salida acumulada se duplica. Si examinamos los datos de las industrias de semiconductores, observaremos que los precios de los circuitos integrados han disminuido en un 28 por ciento por cada duplicación de la experiencia de la industria. Debido al rápido crecimiento de esta joven industria, tales reducciones de costes se han producido a una marcha más acelerada que en las industrias tradicionales; la experiencia de la industria electrónica se ha doblado a un ritmo casi anual. El coste de una función electrónica dada ha estado bajando más rápidamente aún que el coste de los circuitos integrados, ya que la complejidad de los circuitos ha aumentado y bajado el precio de los mismos. Por ejemplo, el coste por bit (dígito binario) de las memorias de acceso aleato-

rio ha descendido a una media del 35 por ciento al año desde 1970, fecha en que se registró el crecimiento más importante en la adopción de las memorias de semiconductores. Los costes bajaron, no sólo debido al tradicional proceso de aprendizaje, sino también a causa de la integración de más bits en cada circuito integrado: en 1970 se produjo un cambio de 256 bits a 1024 bits por circuito, y ahora se está en vías de pasar de 4096 a 16.384 bits por circuito.

Los descensos, por céntuplos, en los precios de los componentes electrónicos desde el desarrollo del circuito integrado son únicos, ya que, si bien otras industrias han presentado curvas de experiencia similares, la industria del circuito integrado no ha tenido par en la duplicación anual de sus salidas a lo largo de bastantes años. Más que servir a un mercado que sólo crece al ritmo del producto nacional bruto o de la población, la industria ha abastecido un fecundo mercado de aplicaciones cada vez más amplias. A medida que cada nueva aplicación consumió más dispositivos microelectrónicos, se ha adquirido más experiencia, comportando mayores reducciones de coste, las cuales a su vez han abierto a los dispositivos mercados más extensos. Antes de la producción de los circuitos integrados en 1960, se fabricaron unos 500 millones de transistores. Considerando que cada transistor representa una función que puede equipararse a una "puerta" lógica o a un bit de memoria en un circuito integrado, el uso anual se ha incrementado 2000 veces, o lo que es lo mismo, se ha duplicado 11 veces en los últimos 17 años. Este asombroso crecimiento favorece las reducciones continuas en los costes.

El principal medio de reducción de costes ha sido el desarrollo de circuitos cada vez más complejos que abaratan el coste por función, tanto al productor de circuitos como al fabricante de equipos. La principal barrera técnica para lograr más funciones por circuito estriba en el rendimiento de la producción. Circuitos más complejos dan lugar a dispositivos mayores y a una probabilidad creciente de defectos, de modo que debe desecharse un porcentaje superior del número total de dispositivos. Cuando el coste de los desechos supera al ahorro obtenido en el montaje subsiguiente y las operaciones de comprobación, el coste por función aumenta en vez de disminuir. El diseño que cumpla mejor la relación coste-eficiencia es un compromiso entre los altos costes de montaje (que aparecen a bajos niveles de integración) y los altos costes de desechos (que aparecen a altos niveles de integración).

Los despliegues tecnológicos se han centrado principalmente en el crecimiento del rendimiento de la producción, a lograr por reducción de la densidad de defectos o por reducción de las dimensiones. Para reducir la densidad de defectos han sido precisas una meticulosa atención al control del proceso y la limpieza del mismo. Basta una partícula de polvo en cualquier proceso crítico para inutilizar un dispositivo; así que la mayor parte de las operaciones deben llevarse a cabo en "locales limpios". Se ha conseguido la reducción de las dimensiones de los elementos básicos, que hace posible reunir mayor número de circuitos complejos dentro de un área determinada, a través de la mejora de la resolución de los procesos de fotograbado. Ahora que se están alcanzando los límites ópticos, ya que las dimensiones de los diseños de circuitos entran en el rango de sólo unas pocas longitudes de onda de luz, se están desarrollando métodos para reducir las dimensiones más



LOS PRECIOS DE LOS CIRCUITOS INTEGRADOS se han ajustado a una curva de experiencia común a muchas industrias, descendiendo un 28 por ciento, aproximadamente, en cada duplicación de la experiencia acumulada por la industria (medida según las unidades producidas). El rápido crecimiento de la industria microelectrónica ha hecho que los índices de reducción de precios sean mayores que en otras ramas.

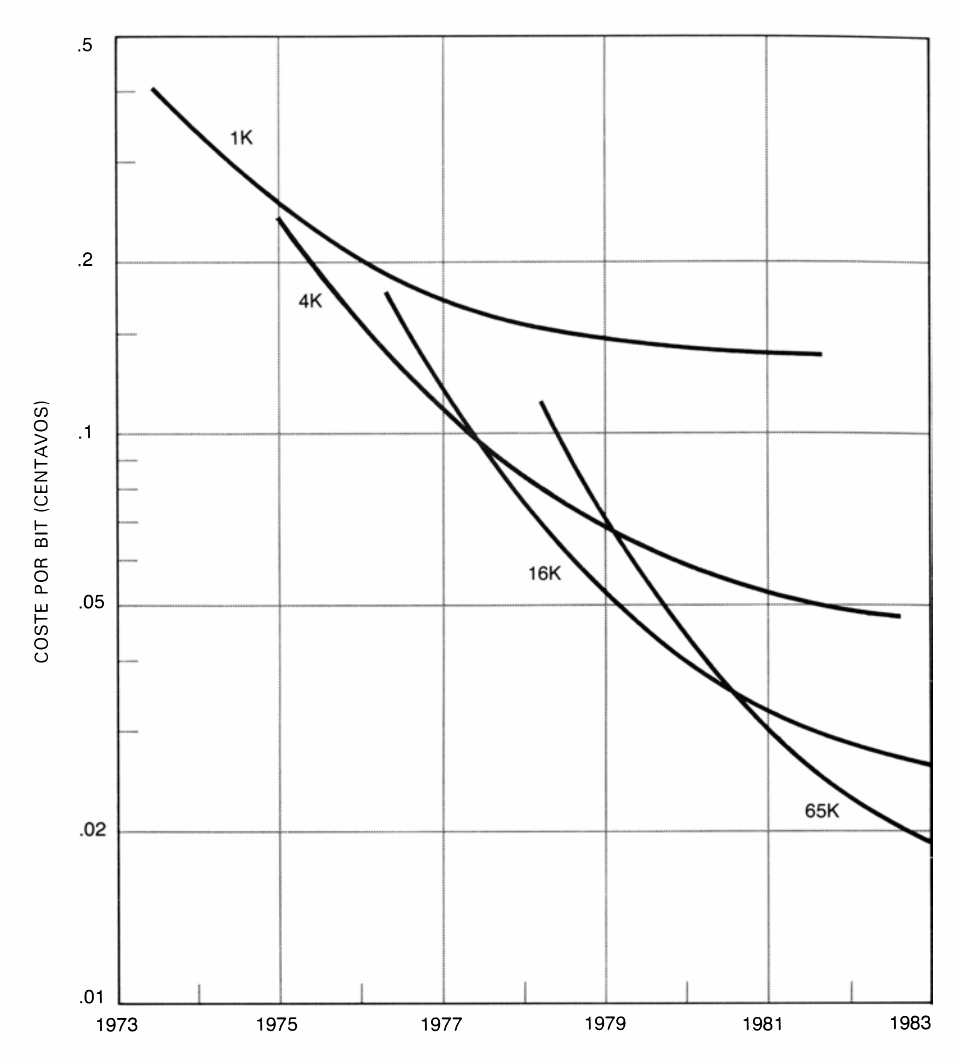
todavía, en los cuales la luz visible se sustituye por haces electrónicos o por rayos X (véase "La fabricación de circuitos microelectrónicos", de William G. Oldham, en este número).

La reducción en el tamaño de los elementos de los circuitos abarata el coste y mejora también el rendimiento básico del dispositivo. Los tiempos de retardo son directamente proporcionales a las dimensiones de los elementos, de modo que el circuito gana en rapidez conforme mengua su talla. De igual modo, la potencia se reduce con el área de los circuitos. Las dimensiones lineales de los elementos pueden reducirse probablemente a la quinta parte del tamaño actual antes de topar con ningún límite fundamental.

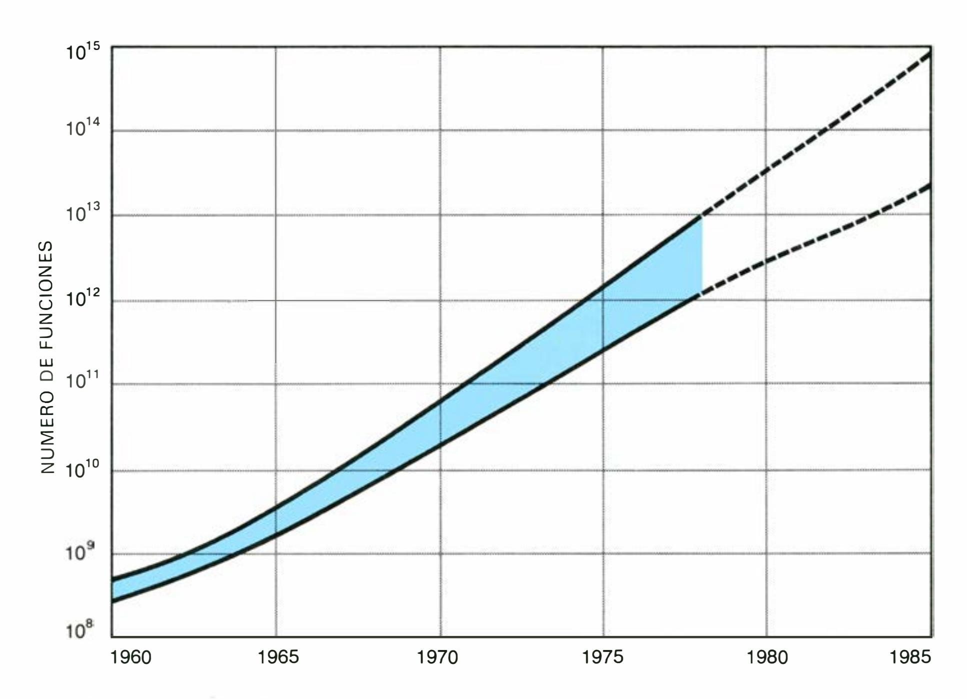
En una industria cuyo producto desciende de precio un 25 por ciento al año, el estímulo para investigar y avanzar es evidentemente intenso. Adelantarse en un año en la presentación de un nuevo producto o de un nuevo proceso puede proporcionar a una compañía un 25 por ciento de ventaja en los precios sobre las compañías rivales; y, a la inversa, un año de retraso pone a una compañía en importante desventaja respecto a sus competidores. El perfeccionamiento del producto es una parte decisiva de la estrategia de las compañías, como cierto es el hecho del envejecimiento de los modelos. La recompensa a una acertada inversión en investigación y desarrollo es enorme, como lo es el castigo al fracaso. Los principales productores de circuitos integrados gastan aproximadamente el 10 por ciento de sus ingresos por ventas en investigación y desarrollo. En un medio de precios constantes se puede decir que la inversión en investigación y desarrollo equivale a una renta anual que proporciona 2,50 dólares por cada dólar invertido. Evidentemente, la mayor parte de esta renta está pagada por los compradores de circuitos integrados o reflejada en las reducciones de precios necesarios para fomentar nuevos mercados.

En este entorno de rápido crecimiento de mercados, súbitos cambios tecnológicos y grandes beneficios para el desarrollo afortunado de nuevos productos o procesos, se han creado y aprovechado gran número de oportunidades empresariales. No deja de ser chocante que aunque los Estados Unidos han encabezado tanto el desarrollo como la comercialización de la nueva tecnología, no fueron las compañías que estaban al frente del negocio de las válvulas de vacío las que procedieron al desarrollo de su sucesor, el transistor. De los 10 principales productores estadounidenses de válvulas en 1955, sólo dos de ellos están entre los 10 primeros productores norteamericanos de semiconductores; cuatro de las 10 primeras compañías de semiconductores se formaron después de 1955, y esas cuatro representan una pequeña fracción de las nuevas y prósperas empresas del ramo. Una y otra vez, el rápido crecimiento del mercado ha hallado a las compañías existentes demasiado ocupadas en intensificar el mercado o series de productos con los que estaban ya comprometidas, como para explorar nuevos mercados o tecnologías más depuradas. De esta suerte quedaba abierta la puerta a nuevas operaciones, típicamente encabezadas (originalmente al menos) por un empresario con una base de conocimientos en investigación o mercado que tuviera suficiente fe en la nueva tecnología y en el nuevo mercado como para arriesgarse. Por fortuna para la economía de los Estados Unidos, había capital disponible a final de los años 50 y en los años 60 para financiar estas aventuras. Se formaron aproximadamente 100 nuevas compañías para producir dispositivos semiconductores; muchas de ellas contribuyeron grandemente al desarrollo de la microelectrónica. En dos de esas aportaciones participé directamente, y fueron: el desarrollo del transistor planar y del circuito integrado planar en Fairchild, cuando esta organización apenas llevaba dos años de funcionamiento, y el desarrollo del microprocesador en Intel Corporation, a los dos años de la fundación de ésta. Hay muchos más ejemplos. El ambiente propicio a las novedades empresariales que reina en los Estados Unidos no tiene parangón en otras naciones industrializadas, y ha sido un factor definitivo del liderazgo norteamericano en este campo.

El crecimiento de la microelectrónica ha creado a su vez otras oportunidades.



EL COSTE POR BIT de las memorias de ordenadores ha descendido y continuará descendiendo, como se muestra aquí, en sucesivas memorias de acceso aleatorio, capaces de manejar desde 1024 (1K) hasta 65536 (65K) bits de memoria. La creciente complejidad de los circuitos que han venido luego sucediéndose es la principal responsable de la reducción de los costes.



LA UTILIZACION ANUAL de funciones electrónicas (transistores, puertas lógicas y bits de memoria) se ha incrementado unas 2000 veces desde que se desarrolló el circuito integrado en 1960. Es de esperar que la utilización aumente en otro factor de 100 durante los próximos diez años.

Ha surgido un abanico de compañías para cubrir las necesidades de los productores de circuitos integrados. Estas compañías suministran todo, desde silicio monocristalino hasta ayudas al diseño controlado por ordenador para equipos de comprobación automática e instrumentos especiales. Frecuentemente, los productos de consumo originales, procedentes de distintos progresos en microelectrónica, fueron en un principio elaborados y vendidos por compañías nuevas. El reloj digital y los juegos de televisión son, entre otros, ejemplos muy familiares.

Cuando el circuito integrado estaba dando sus primeros pasos, Patrick E. Haggerty, de Texas Instruments, llamó la atención sobre la creciente penetración de la electrónica; predijo que las técnicas electrónicas seguirían desplazando a otros métodos de control, alcanzando a casi todos los aspectos de nuestra vida. Tal desplazamiento ha ocurrido principalmente porque la industria electrónica ha podido fabricar elementos funcionales cada vez más complejos a precios siempre decrecientes. Los componentes mecánicos de la calculadora y del reloj han sido substituidos por circuitos integrados, que resultan menos caros y ofrecen mayor flexibilidad. Estamos asistiendo a la substitución de las funciones electromecánicas de las máquinas expendedoras, los semáforos y otros. En un futuro próximo, el automóvil será controlado por un ordenador con la consiguiente mejora en rendimiento y reducción de contaminantes. Todas estas aplicaciones son puras extensiones de las aplicaciones tradicionales de la electrónica al manejo de información en mediciones, comunicación y manipulación de datos. Se ha dicho con frecuencia que así como la revolución industrial capacitó al hombre para aplicar y controlar potencias físicas mayores que las que sus propios músculos pudieran desarrollar, la electrónica ha prolongado su poder intelectual. Pues bien, la microelectrónica lleva ese poder mucho más lejos.

Para 1986, el número de funciones electrónicas incorporadas cada año a una amplia gama de productos puede esperarse que aumente unas 100 veces respecto al número existente hoy. La curva de experiencia prevé que el coste por función habrá descendido, para entonces, hasta una veinteava parte del coste de 1976, es decir, una reducción del 25 por ciento al año. Con tales precios, los dispositivos electrónicos serán aprovechados en vastas zonas, potenciando el servicio de correos, extendiendo las bibliotecas y haciendo sus contenidos más accesibles, proporcionando entretenimiento, propagando los conocimientos para fines educacionales y realizando muchas tareas rutinarias del hogar y la oficina. En la proliferación exponencial de productos y servicios dependientes de la microelectrónica es en donde se manifestará la verdadera revolución microelectrónica.

Elementos del circuito microelectrónico

La unidad funcional básica de un circuito electrónico moderno es el transistor, el cual, con el advenimiento de la tecnología microelectrónica, ha abaratado los costes de tal manera que su empleo se presenta en miles de aplicaciones

James D. Meindl

Los dispositivos electrónicos están compuestos por los elementos "activos" de los circuitos, como los transistores, que se encuentran combinados con otros elementos, cuales son las resistencias, los condensadores y los inductores. Tiempo atrás, el sistema de fabricación de los circuitos electrónicos consistía en producir cada elemento por separado, para luego cablearlos juntos mediante conductores metálicos y así formar los dispositivos propiamente dichos.

El advenimiento de los circuitos microelectrónicos no ha cambiado substancialmente la naturaleza de las unidades funcionales básicas: los circuitos microelectrónicos constan, también, de transistores, resistencias, condensadores y componentes similares. La gran diferencia existente entre ambos radica en que todos estos elementos, y sus interconexiones respectivas, se fabrican ahora sobre un substrato único en una serie única de operaciones.

La materia prima de los circuitos microelectrónicos es el silicio; y así, el desarrollo de la microelectrónica se debe en gran parte al descubrimiento de técnicas para fabricar las distintas unidades funcionales en (o sobre) un cristal de este material semiconductor. Se han desarrollado métodos para fabricar la mayoría de los dispositivos normalizados, pese a que los inductores no pueden emplearse de forma conveniente. No sólo se adoptaron y miniaturizaron los diseños de la antigua tecnología. Con el cambio experimentado en el tamaño, se ha logrado un cambio importante en los recursos disponibles por el diseñador, seguido por un cambio en la forma en que tales recursos se organizan. Los diseñadores han asignado un número cada vez mayor de funciones a los elementos del circuito que resultan de fabricación fácil en silicio y que, además, prestan un mayor rendimiento; tal es el caso del transistor, del que se han desarrollado varios tipos y, con ellos, toda una familia de elementos asociados.

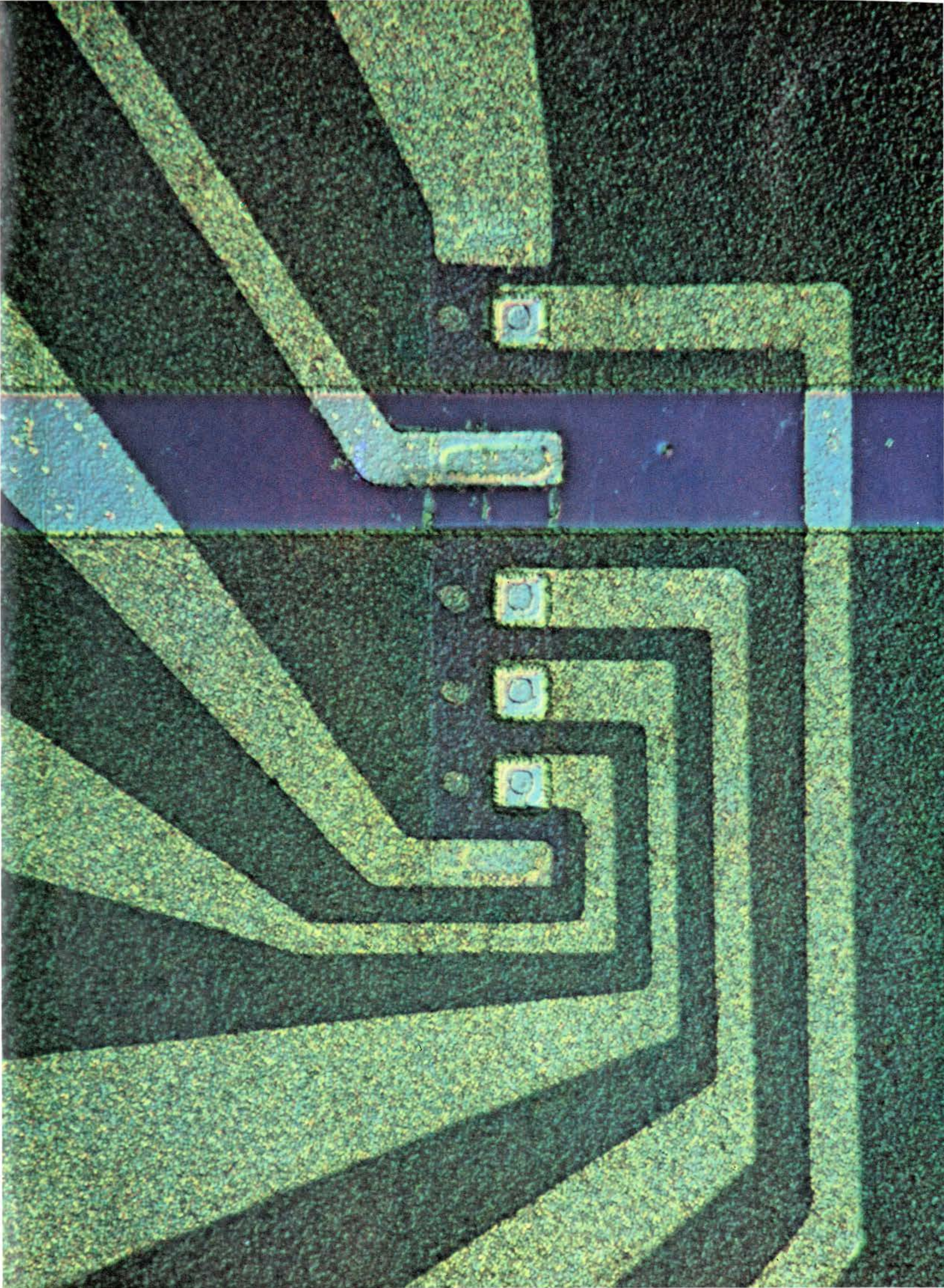
Se puede distinguir entre los elementos activos del circuito, representados por el transistor, y el resto de los dispositivos electrónicos que, por contraste, se denominan elementos pasivos. El elemento activo puede cambiar su estado en respuesta a una señal externa, mientras que las resistencias, condensadores e inductores no pueden hacerlo.

Cualquier material portador de corriente eléctrica, incluidos los cables, muestra todas las características de los elementos pasivos, a saber: resistencia, capacitancia e inductancia. Los componentes útiles poseen una de ellas de forma más destacada. La resistencia eléctrica de un cable metálico puede interpretarse como la perturbación de un movimiento ordenado de los electrones, causada por determinadas interacciones en la estructura atómica del material. Los elementos resistivos, construidos como componentes discretos, emplean un medio que presenta una notable resistencia por unidad de longitud, como el carbón o una aleación de nicromo.

La capacitancia y la inductancia son efectos que pueden atribuirse a los campos electromagnéticos generados por una corriente o carga eléctrica. La capacitancia es una medida del campo eléctrico que circunda a un conductor. La capacitancia obtiene los valores más elevados cuando extensas áreas de conductores que portan cargas eléctricas opuestas se aproximan una a otra; de ahí que los condensadores discretos estén fabricados, frecuentemente, por placas o panes metálicos separados por una capa fina aislante. La inductancia representa la energía almacenada en el campo magnético originado por una corriente eléctrica. A fin de concentrar este campo, los inductores se fabrican arrollando un conductor en una bobina, cuyo núcleo suele ser de un material ferromagnético.

En los aparatos microelectrónicos, el carbón o el nicromo de una resistencia, así como la intercalación de conductores de un condensador o los arrollamientos de un inductor, no están disponibles. Todos los componentes del circuito deben fabricarse en cristal de silicio o bien sobre la superficie de éste. El silicio dista mucho de ser un material idóneo para estas funciones, ya que sólo pueden conseguirse valores de resistencia y capacitancia muy modestos. Y no hay forma de construir inductores microelectrónicos prácticos. Por otra parte, el silicio es un material inigualable para fabricar

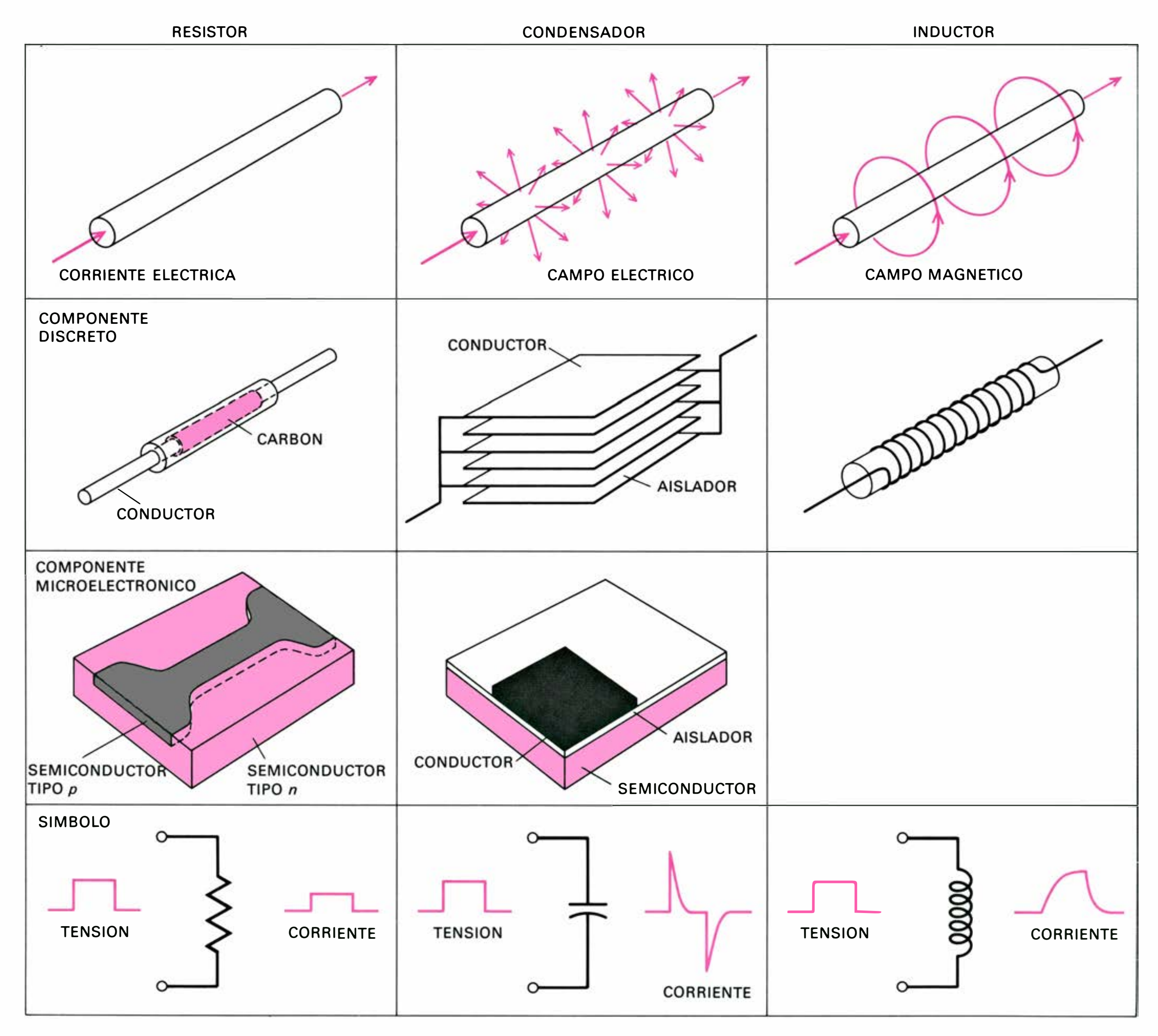
PAR DE CIRCUITOS LOGICOS, en la microfotografía de la página siguiente, fabricados formando una unidad en la superficie de una pastilla de silicio. Los circuitos constan de varios transistores interconectados, que son, con mucho, los elementos de circuito más importantes en punto a dispositivos electrónicos. De los cuatro conductores que circunscriben el dispositivo y lo abordan desde la derecha, el de la parte superior origina el contacto con una región de un transistor cuya corriente fluye paralela a la superficie de la pastilla, y surge a través del conductor que termina en la barra horizontal de color azul brillante. Debajo de este conductor se encuentran otros tres transistores de un tipo distinto: en ellos, la corriente fluye desde el substrato hacia la fila de contactos, que está situada debajo de la barra azul. El circuito es representativo de la tecnología denominada lógica integrada de inyección, o bien, I^2L. Un rasgo característico de los circuitos I^2L es que algunas regiones de la pastilla operan como elementos de más de un transistor. En la barra azul puede apreciarse la superficie de la pastilla. La textura abrupta corresponde al silicio policristalino. Las variaciones en el color no son intrínsecas a la naturaleza del silicio, sino que se deben a interferencias registradas en las capas de dióxido de silicio que recubren la superficie. El dispositivo lo hizo en el laboratorio del autor, en la universidad de Stanford, Roderick D. Davies.



transistores y la abundancia de estos componentes activos en microelectrónica compensa con creces las deficiencias presentadas en los elementos pasivos.

Una de las propiedades que definen a los transistores y que los hace indispensables en microelectrónica es la capacidad de ganancia o amplificación; esta propiedad debe entenderse considerando los elementos activos y pasivos como registros de vuelo o "cajas negras" cuya tarea interna carece de interés, pero cuyo comportamiento puede ser examinado en los terminales de entrada y salida. Una señal aplicada a los terminales de entrada de una caja negra que contenga una resistencia, un condensador o un inductor, puede transformarse de muchas formas, pero invariablemente la señal aplicada quedará reducida en su potencia. Pero, si disponemos de otra caja que contenga un transistor, se puede transformar una señal de potencia baja en otra de potencia alta.

Para comprender por qué los transistores y otros elementos del circuito pueden fabricarse con silicio, conviene considerar la naturaleza física de los materiales semiconductores. En un conductor, un metal por ejemplo, la corriente se transporta por electrones que pueden moverse libremente en el enrejado atómico de la substancia conductora; por contra, en un aislador, todos los electrones están fuertemente ligados a los átomos o moléculas, de suerte que ninguno puede actuar como portador de carga eléctrica. En los semiconductores, la situación existente es intermedia entre las dos an-



LOS ELEMENTOS PASIVOS DE UN CIRCUITO electrónico se hallan compuestos por resistores, condensadores e inductores. La resistencia es la energía disipada por los electrones en su movimiento a través de la estructura atómica de un conductor; la capacitancia mide de la energía almacenada en el campo eléctrico que rodea un conductor cargado; la inductancia mide la energía almacenada en el campo magnético originado por la corriente eléctrica. Los elementos resistivos construidos como componentes discretos están formados por carbón u otra substancia que sea mal conductor de la electricidad, como la aleación de níquel-cromo. Un resistor que forme parte de un circuito microelectrónico consta de una cinta delgada de material semiconductor, circundada por otro semiconductor de tipo contrario. Los condensadores discretos están formados por capas aislantes y conductoras intercaladas; por su parte, el condensador microelectrónico está hecho de una fina capa de aislante seguida por una capa metálica situadas en la superficie de un cristal semiconductor. De este modo sólo pueden obtenerse valores de capacitancia pequeños. El inductor discreto está constituido por un hilo bobinado sobre un núcleo de material ferromagnético; todavía no se han encontrado métodos de fabricación satisfactorios para los inductores de uso microelectrónico. Los símbolos de cada elemento pasivo, y la señal de corriente que fluye por ellos cuando se les aplica una tensión, aparecen en la parte inferior.

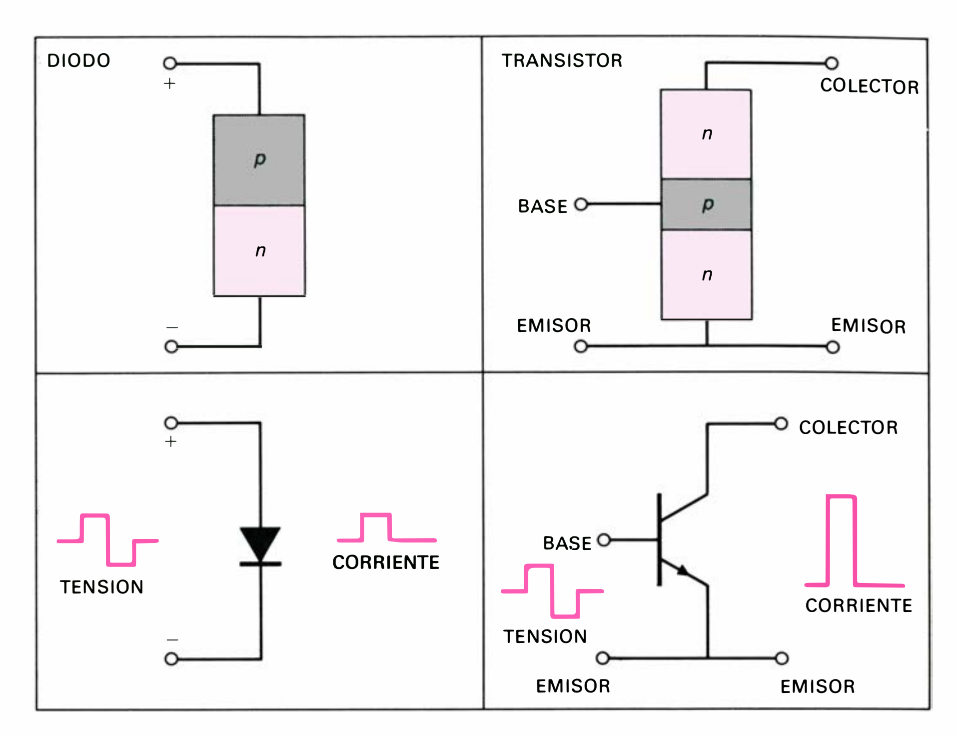
teriormente reseñadas; por ello, los portadores libres de carga no siempre están presentes, pero sí pueden generarse con un leve consumo de energía.

El átomo de silicio tiene cuatro electrones de valencia en su envoltura externa. Ahora bien, en el silicio sólido, parejas de estos electrones, compartidos por átomos vecinos, están dispuestos simétricamente, de forma que cada uno está rodeado por ocho electrones compartidos. Al hallarse sometidos todos los electrones a las fuerzas de enlace entre átomos, queda explicado el hecho de que un cristal de silicio puro sea mal conductor de electricidad.

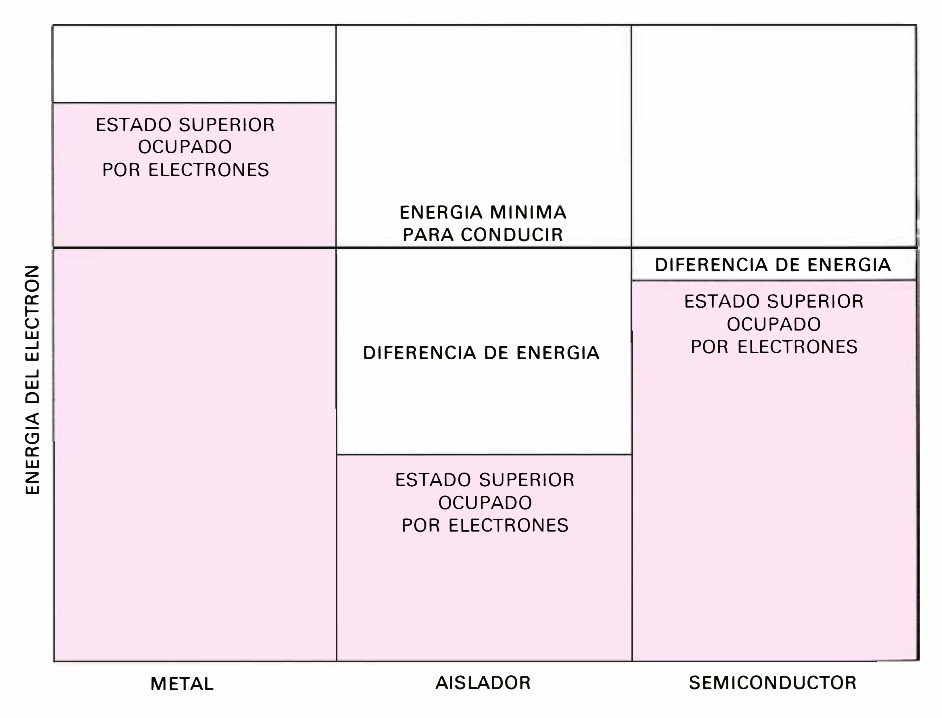
Los dispositivos semiconductores se obtienen mediante la introducción en el cristal de un número controlado de átomos de impurezas, por lo que el proceso recibe el nombre de impurificado. Supongamos que parte de un cristal de silicio se impurifica con fósforo, que es un elemento cuyos átomos tienen cinco electrones de valencia en su envoltura externa; el átomo de fósforo puede desplazar a otro átomo de silicio sin romper la estructura del cristal, si bien el electrón de más que comporta no tiene sitio en los enlaces interatómicos. En ausencia de estímulo externo, el electrón supernumerario permanece en las proximidades del átomo de impureza, pudiendo movilizarse en caso de aplicársele una pequeña tensión al cristal.

El silicio puede impurificarse también con boro, cuyos átomos tienen tres electrones de valencia en su capa externa; cada átomo de boro insertado en la estructura del silicio origina la deficiencia de un electrón; a la que se le denomina hueco. En circunstancias normales, el hueco permanece asociado a un átomo de impureza, pero puede iniciar el movimiento en respuesta a la tensión aplicada al cristal. El hueco no es ninguna partícula, sino la ausencia de un electrón en una posición en donde éste podría encontrarse en una estructura pura de átomos de silicio. A pesar de ello, el hueco presenta carga eléctrica positiva y puede conducir corriente eléctrica. El hueco se mueve por la estructura cristalina como si fuera una burbuja en un líquido. Un átomo adyacente transfiere un electrón al átomo de impureza, "llenando" el hueco y creando otro nuevo en su propia nube electrónica; el proceso vuelve a repetirse, de forma que el hueco pasa de átomo en átomo.

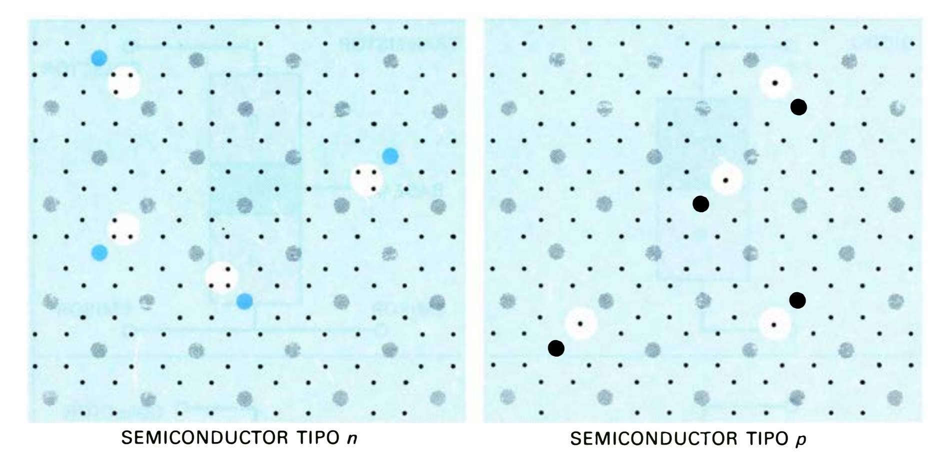
El silicio impurificado con fósforo u otro elemento pentavalente es semiconductor de tipo *n*, en donde *n* representa la carga eléctrica negativa resultante de



DISPOSITIVOS SEMICONDUCTORES, a partir de los cuales se construyen los circuitos microelectrónicos. Comprenden el diodo *(izquierda)* y el transistor. Uno y otro dispositivos están formados por regiones yuxtapuestas de silicio impurificado con elementos contaminantes; los dos tipos de silicio impurificado reciben el nombre, respectivamente, de tipo *n* y tipo *p*. La asimetría constituye la propiedad esencial del diodo. Tal como se muestra en la ilustración, el diodo transmite una señal positiva y bloquea otra de polaridad negativa. En ese aspecto, el transistor se comporta también asimétricamente, si bien presenta una propiedad adicional importante. Un transistor es capaz de amplificar una señal de baja potencia gracias a una fuente de alimentación externa.



LA ESTRUCTURA ELECTRONICA de metales, aisladores y semiconductores determina sus propiedades eléctricas. Cada uno de los electrones que se encuentran en una substancia ocupa un estado único y discreto; los estados habilitados se llenan por orden, comenzando por el de energía más baja. La conducción requiere electrones con una energía relativamente alta (por razones de simplificación, la energía mínima para la conducción aparece en la figura como si fuera la misma para todos los materiales, aunque de hecho existen substanciales diferencias). En un metal, el estado superior ocupado por electrones tiene una energía que supera el mínimo de conducción, por lo que siempre se hallan presentes electrones de conducción; se desplazan libremente por entre el enrejado atómico del metal. En los aislantes existe una larga distancia entre la energía máxima de los electrones y la que se requiere para conducir; esta diferencia se interpreta como la energía necesaria para liberar un electrón del átomo al que va ligado. También existe diferencia en la estructura electrónica de un semiconductor aunque es pequeña. Cuando no sufren alteraciones, los electrones de un semiconductor permanecen en las proximidades de átomos particulares, si bien pueden liberarse mediante un pequeño gasto energético.



LA IMPURIFICACION DEL SILICIO por átomos de impureza altera la estructura electrónica de la substancia, de suerte que los portadores de carga eléctrica se sueltan con facilidad del enrejado atómico. Un átomo de silicio *(gris)* posee cuatro electrones en su valencia o capa externa; en un cristal puro tales electrones forman parejas que comparten con los átomos adyacentes. En consecuencia, cada átomo se encuentra rodeado por ocho electrones, que es una configuración naturalmente estable. Substituyendo unos cuantos átomos de silicio por los de otro elemento, el fósforo por ejemplo *(color)*, que tiene cinco electrones en su capa de valencia, se obtiene un semiconductor de tipo *n*. El electrón sobrante no puede participar en los enlaces entre los átomos del cristal, pasando a ser de esta manera un portador móvil de carga. En un semiconductor de tipo *p*, la impureza introducida es un elemento, como el boro *(negro)*, que posee tres electrones en su valencia o capa externa. Cada átomo de impureza origina la deficiencia de un electrón, denominada hueco. Cada uno de éstos está dotado de una carga neta positiva; al someterlo a una tensión, el hueco se desplaza de átomo en átomo a través de la estructura cristalina. La simbología *n* y *p* hace referencia a las polaridades negativa y positiva, respectivamente, de los portadores de carga. En la ilustración presente se ha exagerado notablemente la concentración de átomos de impureza.

la conducción de electrones. Si el impurificado se efectúa con boro u otro elemento trivalente, se origina un semiconductor de tipo *p*, cuya letra hace referencia a que en sus huecos hay presencia de cargas positivas.

El dispositivo semiconductor más simple es el diodo que resulta de la unión de regiones *n* y *p*, en un cristal único de silicio. Cuando se aplica una tensión positiva a la región *p* y otra negativa a la región *n*, se establecen corrientes contrarias de electrones y huecos. Los huecos se localizan en la región *p* y son repelidos por la carga positiva aplicada al terminal *p* y atraídos hacia el terminal *n*, de forma que se origina un flujo electrónico a través de la unión. Por su parte, los electrones de la región *n* salen repelidos en dirección opuesta. La fuerte corriente que se origina entonces recibe el nombre de corriente directa de diodo.

Si se invierten las conexiones del diodo, los huecos tenderán hacia el terminal de la región *p*, que tiene ahora carga negativa, y los electrones retrocederán hacia la región *n*, hacia el terminal positivo. No existe flujo de corriente a través de la unión. En realidad, siempre se observa una corriente "inversa" muy pequeña, originada por algunos electrones encontrados en el silicio de tipo *p* y por huecos en el de tipo *n*. Tales portadores de corriente "minoritarios" siempre se hallan presentes, aunque su concentración es muy baja.

El diodo no está capacitado para efectuar amplificación; por tanto, no sirve como elemento activo; además es poseedor de una propiedad que lo diferencia de todos los elementos restantes de tipo pasivo. Los dispositivos resistivos, los condensadores e inductores son elementos simétricos; es decir, sus efectos sobre una señal son idénticos cualquiera que sea la polaridad de la señal y la manera como estén conectados en el circuito. La propiedad más sobresaliente del diodo radica en su asimetría; es decir, presenta una baja resistencia a señales de polaridad determinada, mientras que el valor resistivo es elevado para la misma señal dotada de polaridad opuesta.

Si añadimos una tercera región impurificada a un diodo, obtenemos el transistor, de suerte que se dé, por ejemplo, un emparedado de la región *p* entre dos *n*. Una de las áreas *n* se denomina emisor y la otra colector; la región *p*, existente entre ambos, recibe el nombre de base. En razón de su estructura, el transistor puede considerarse como la unión de dos diodos en un solo cristal de silicio. Tal como cabe esperar, el funcionamiento del transistor está condicionado por las tensiones relativas que se apliquen a las tres regiones anteriormente definidas.

Supongamos que al emisor de tipo *n* se le aplica un potencial de cero voltios, a la base *p* una pequeña tensión positiva y al colector de tipo *n* otra tensión positiva mayor. Con tal disposición de tensiones, tendremos que el emisor y la base representan un diodo que opera en forma directa, ya que los huecos son arrastrados hacia el emisor y los electrones dirigidos a la base. En contraste con ello, las tensiones de base y colector presentan polaridades inadecuadas para la conducción; por tanto, sólo una pequeña corriente inversa de valor despreciable fluye a través de la unión. Pese a ello, el colector no permanece inactivo, puesto que los electrones impulsados por el emisor hacia la base se ven transportados hacia el colector en virtud de un proceso de difusión similar a la mezcla de dos gases. En un transistor de buen diseño, casi todos los electrones inyectados se difunden a través de la base, y se origina un considerable flujo de corriente de emisor a colector.

Si la tensión positiva de la base se reduce a cero, o si se conecta un potencial negativo a la base, la corriente directa del diodo del emisor a la base se suspende; además, al no ser impulsados los electrones a la base, cesa la corriente de emisor a colector, ya que la corriente de colector está controlada por la corriente de base. Variando la polaridad de la tensión aplicada a la base, puede activarse o desactivarse la corriente de colector; entre estos dos estados, la corriente de colector es aproximadamente proporcional a la corriente de base. Si las regiones impurificadas de un transistor presentan las dimensiones y la disposición correctas, la corriente de colector puede ser mucho mayor que la corriente de base, por lo que cabe decir que el transistor presenta una ganancia. Normalmente, la relación entre las corrientes de base y de colector alcanza un valor 100, pudiendo conseguirse amplificaciones con valores superiores a 1000.

El apremio dimensional más importante en el diseño de tal transistor radica en que la base esté tan ajustada que virtualmente todos los electrones emitidos por el emisor sean conducidos al colector a través de la base; además, una base así ajustada reduce el tiempo de tránsito de un electrón e incrementa consiguientemente la velocidad a la que el dispositivo cambia de estado.

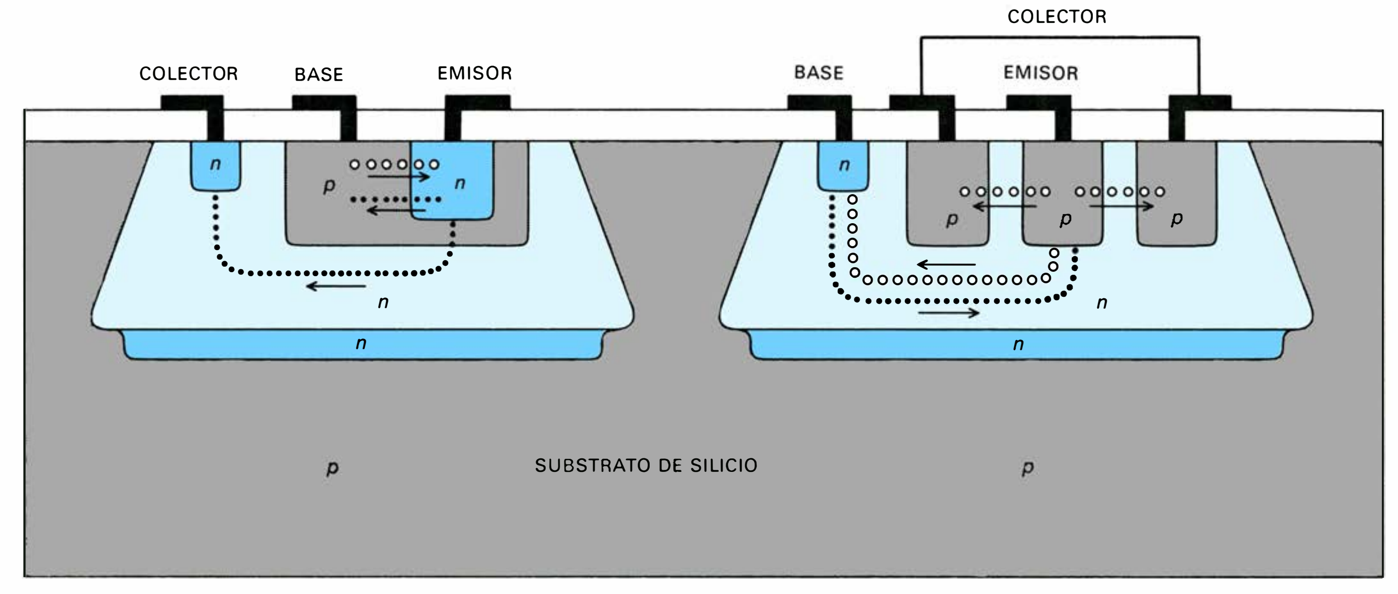
El transistor anteriormente reseñado se denomina *npn;* este término denota la secuencia de regiones impurificadas que se encuentran en el cristal de silicio. Por lo que respecta al transistor tipo *pnp*, que es el dispositivo complementario, cabe decir que está constituido por la inserción de una base tipo *n* entre re-

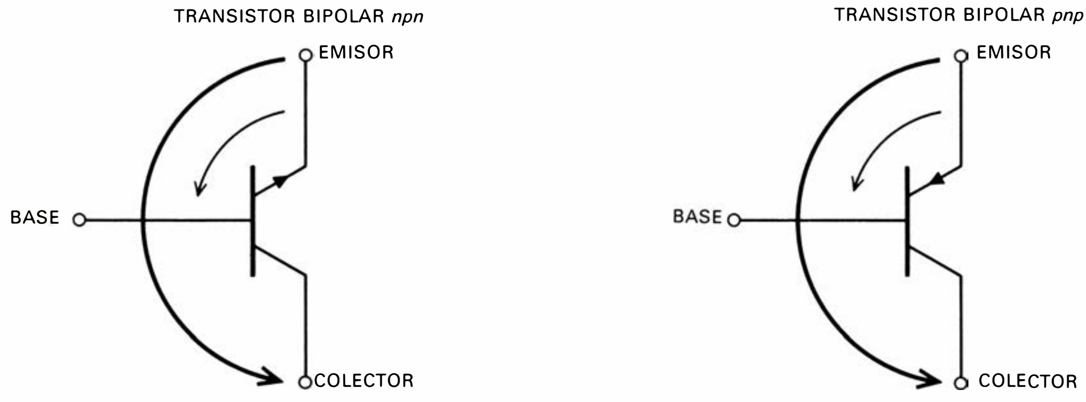
giones tipo *p* que definen al colector y emisor. Los principios de funcionamiento son iguales que en el transistor *npn*, pero con todas las polaridades invertidas, de forma que, para obtener un funcionamiento normal, la base y el colector deben ser negativos respecto al emisor. Las polaridades de los portadores de carga también están invertidas, ya que la corriente del emisor al colector consiste en la introducción de huecos.

Los transistores *npn* y *pnp* forman la clase de los denominados transistores de unión o bipolares, debido a que los portadores de carga poseen ambas polaridades. Fueron descubiertos en 1948 por John Bardeen, Walter H. Brattain y William Shockley, de los Bell Telephone Laboratories.

Existe una segunda clase fundamental de transistores denominada transistores de efecto de campo. Aunque concebidos casi 25 años antes que los transistores bipolares, su fabricación a gran escala no se llevó a cabo hasta la década de los 60. De entre los diferentes tipos existentes, hay uno que es de uso común en la microelectrónica: el transistor de efecto de campo de óxido metálico semiconductor. Su denominación remite a los materiales empleados en su fabricación, y se conoce bajo las siglas MOSFET, del inglés "metal-oxide-semiconductor field-effect transistor".

En un MOSFET típico, se encuentran dos islas de silicio de tipo *n* en un substrato de tipo *p*, efectuándose las conexiones directamente sobre las islas, las cuales reciben el nombre de surtidor y drenador respectivamente. Sobre la superficie del silicio y en el canal existente entre el surtidor y el drenador se forma una capa de dióxido de silicio (SiO_2); sobre la capa del óxido se deposita otra capa metálica, originándose la formación de un tercer electrodo denominado puerta. El dióxido de silicio presenta unas características aislantes excelentes, de modo que la puerta no tiene conexión directa con el substrato semiconductor. Sin embargo, la puerta está acoplada al





DIOXIDO DE SILICIO (SIO_2)

SILICIO TIPO *p*

SILICIO TIPO *n*

SILICIO TIPO *n* FUERTEMENTE IMPURIFICADO

CONDUCTOR DE ALUMINIO

••••••• ELECTRONES

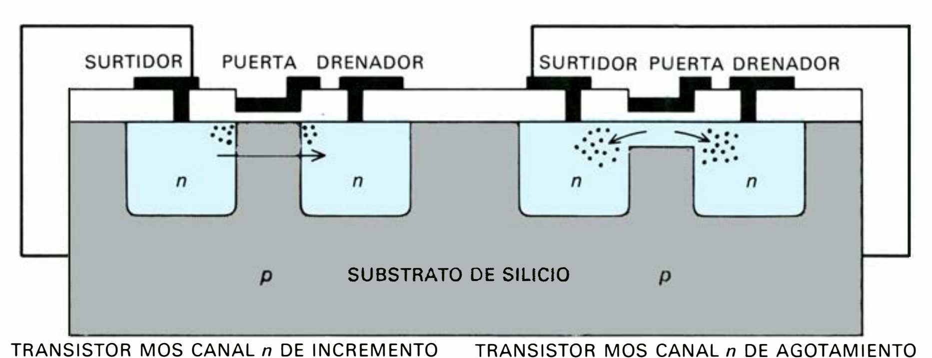
oooo HUECOS

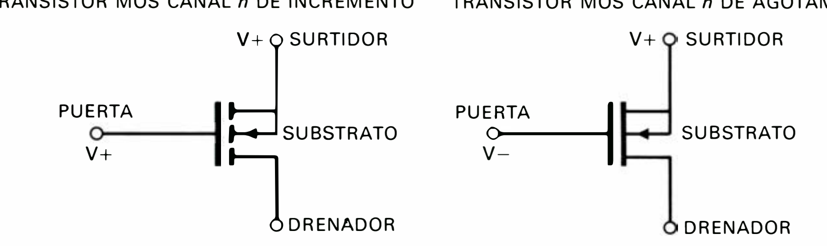
TRANSISTORES MICROELECTRONICOS, fabricados en un cristal único de silicio a través de una serie de operaciones permitidas únicamente en la superficie de la pastilla. En la ilustración se aprecia cómo toda la pastilla queda contaminada por impureza de tipo *p*; luego se forman las islas de silicio de tipo *n*. Inmediatamente se crean áreas menores de tipo *p* y de tipo *n* en el interior de esas islas a fin de determinar los tres elementos fundamentales del transistor, a saber: el emisor, la base y el colector. En un transistor de tipo *npn (izquierda)*, se aplica una tensión positiva a la base y al colector, a resultas de lo cual los huecos se desplazan desde la base hasta el emisor y se facilita el paso de los electrones desde el emisor hasta la base. Son muchos los electrones que emigran de la base para alcanzar el colector, resultando por tanto la corriente emisor-colector mayor que la corriente emisor-base. El dispositivo presenta ganancia, porque si se aplica una pequeña señal a la base se puede controlar otra mayor en el colector. Por sencillez en la fabricación, un transistor *pnp (derecha)* suele construirse de un modo distinto, según el cual la corriente emisor-colector es lateral, en vez de serlo vertical. Los principios de funcionamiento siguen siendo los mismos, con la salvedad de que todas las polaridades son inversas. Las conexiones con los elementos del circuito se realizan mediante conductores de aluminio depositados sobre una capa aislante de dióxido de silicio; algunas zonas de silicio de tipo *n* se hallan fuertemente impurificadas a fin de mejorar su conductividad. Las islas de tipo *n* se emplean para lograr el aislamiento de los transistores. Por cuanto los portadores de carga de ambas polaridades participan en la operación de tales dispositivos en cuestión, éstos reciben el nombre de transistores bipolares.

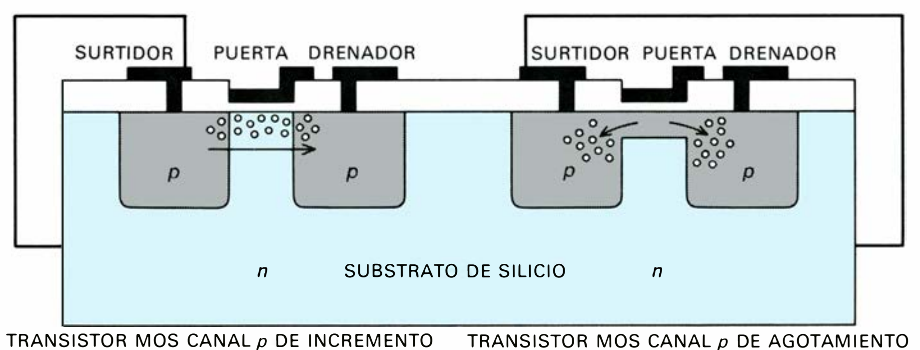
silicio por capacitancia; es decir, el campo eléctrico generado por cualquier carga aplicada en el electrodo de la puerta puede condicionar el movimiento de los portadores de carga en el canal semiconductor.

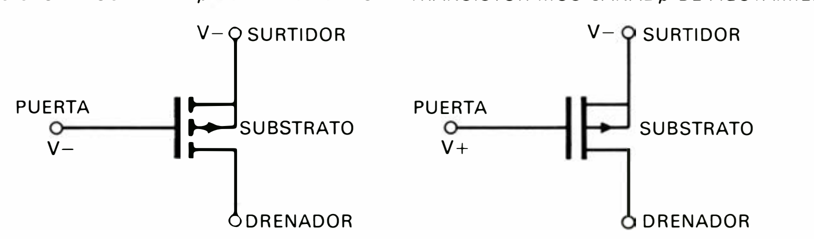
En funcionamiento normal, surtidor y substrato están conectados por un conductor externo, a un potencial de cero voltios. Al drenador se le aplica una tensión de valor positivo. No existe flujo electrónico entre surtidor y substrato porque ambos permanecen conectados a tierra; entre el drenador y el substrato aparece una pequeña corriente inversa, como ocurre en el diodo. Cuando el transistor está en condición estática (sin aplicar tensión a la puerta), el canal de tipo *p* situado debajo de la puerta contiene una mayoría de huecos, y son pocos los electrones que pueden ser atraídos hacia el potencial positivo del drenador. Si se aplica a la puerta una tensión positiva, el campo eléctrico atraerá una mayoría de electrones hacia una capa fina situada en la superficie del cristal y próxima a la puerta; entonces, debido a la presencia de numerosos electrones en una región que normalmente es de tipo *p*, la superficie recibe el nombre de invertida. Esta inversión origina un canal continuo de tipo *n* entre el surtidor y el drenador, con lo que se produce un gran paso de corriente. Al igual que los transistores bipolares, los MOSFET pueden actuar como amplificadores, aunque la ganancia suele especificarse como relación de tensión en vez de relación de corriente.

Conviene observar que, mientras que en la corriente de base de un transistor bipolar participan electrones y huecos, sólo se presenta un tipo básico de portadores de carga en el canal invertido de un MOSFET. En el elemento que he descrito estos portadores son electrones y los transistores se denominan MOSFET de canal *n*, o simplemente transistor MOS *n*. También puede construirse otro dispositivo complementario al anterior, el cual conste de dos islas de tipo *p* en un substrato de tipo *n* con igual acoplo de puerta por capacitancia sobre el canal. En este caso se observa que todas las polaridades son inversas y que los portadores de carga son huecos, y no electrones, por lo que reciben el nombre de transistores MOS *p*.









LOS TRANSISTORES DE EFECTO DE CAMPO, fabricados según la técnica del óxido metálico semiconductor (MOS), difieren de los transistores bipolares en que sólo un tipo de portador de carga actúa en cada dispositivo. Los transistores que emplean electrones se denominan transistores MOS de canal *n*; los que emplean huecos, transistores MOS de canal *p*. En un dispositivo MOS *n (arriba)* se forman dos islas de silicio de tipo *n*, denominadas surtidor y drenador, en un substrato de tipo *p*. Sobre el canal existente entre surtidor y drenador está localizado un electrodo metálico, la puerta, que, mediante una fina capa de dióxido de silicio, queda protegida del contacto con el material semiconductor. En los transistores incrementados *(arriba, izquierda)*, la tensión positiva que se aplique al drenador ejerce una fuerza de atracción sobre los electrones disponibles del surtidor, pero aquéllos no pueden pasar por el canal de tipo *p*, que es rico en huecos. Sin embargo, al aplicar una carga positiva a la puerta, el campo eléctrico que se origina atrae a los electrones hacia una capa fina situada en la superficie del canal, creándose un flujo de electrones desde el surtidor hasta el drenador. En los transistores MOS *n* de agotamiento *(arriba, derecha)* existe un canal continuo de silicio de tipo *n* entre el surtidor y el drenador, de modo que el transistor normalmente conduzca; sólo cuando se aplica una tensión negativa a la puerta se detiene la corriente, porque entonces los electrones son expulsados del canal. Formando islas de material de tipo *p* en un substrato de tipo *n*, se pueden construir los correspondientes dispositivos MOS *p* *(abajo)*; aquí los portadores de carga son los huecos. Puesto que los transistores MOS no necesitan islas, pueden encapsularse en una pastilla con una densidad mayor que los transistores bipolares.

Un último tipo de MOSFET se fabrica conectando dos islas de tipo *n* a un canal *n* continuo, situado bajo el condensador de puerta. Dada la presencia de un canal conductor, cuando la puerta no esté excitada se permitirá el paso de corriente de surtidor a drenador; por otra parte, si se aplica un potencial negativo a la puerta, ésta desalojará a los electrones del canal, detendrá el paso de corriente y pondrá el transistor en corte, por lo que se le denomina transistor de efecto de campo MOS de agotamiento. Puede construirse otro dispositivo de igual tipo con canal *p*. Los transistores cuya conducción se activa aplicando tensión a la puerta se conocen como dispositivos de incremento. Resumiendo todo lo anteriormente expuesto, podemos decir que existen cuatro tipos de transistores MOS, a saber: de canal *n*,

de canal *p* y, en ambos casos, de incremento y de agotamiento.

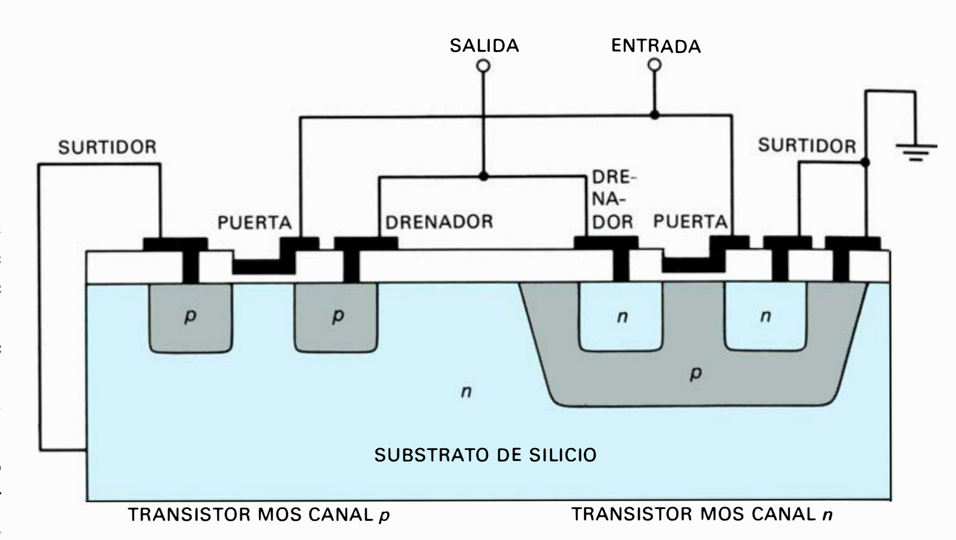
Las dimensiones tienen tanta importancia en la fabricación de los transistores MOS como en la de los bipolares. Para un MOSFET, las dimensiones críticas se refieren al grosor de la capa de óxido situada debajo del electrodo de la puerta y a la distancia existente entre el surtidor y el drenador. Cuando se aplica tensión a la puerta, la sensibilidad de la respuesta del transistor es inversamente proporcional al espesor de la capa metálica.

La materia prima de un dispositivo microelectrónico está constituida por una pastilla de silicio impurificado con distintas concentraciones de elementos de tipo *p* y *n*. Están, asimismo, los conductores, que pueden ser de metal (aluminio) o de silicio policristalino fuertemente impurificado, el cual presenta un alto grado de conductividad. Por último, el dióxido de silicio muestra un óptimo poder aislante.

La fabricación del circuito debe efectuarse enteramente desde la superficie. Las áreas sometidas al proceso de impurificado se definen mediante procedimientos fotolitográficos, tras de lo cual se permite que el elemento que se destina a la impurificación se difunda por la estructura del silicio, debiendo repetirse la operación varias veces a fin de conseguir el impurificado de las regiones que se desean, las cuales con relativa frecuencia se encajan unas en otras. Las áreas de óxido se depositan de forma similar por un proceso de calentamiento de la pastilla en presencia de oxígeno. Los conductores se depositan sobre una gruesa capa de óxido que cubre la pastilla, abriéndose paso a través de la barrera aislante sólo cuando sea necesario efectuar el contacto eléctrico con el silicio.

Mediante estas técnicas puede fabricarse un número limitado de elementos pasivos del circuito. Las resistencias, o dispositivos resistivos, se forman por la circunscripción de una tira fina de silicio de tipo *p* dentro de una isla de material de tipo *n*. La corriente que circula a través del componente queda confinada en la tira de silicio manteniéndola a una tensión negativa con respecto a la isla de tipo *n*, por lo que solamente puede circular corriente inversa a través de la unión *pn*.

Resultaría una elección inverosímil optar por el silicio impurificado como material resistivo para la fabricación de componentes discretos, ya que no presenta suficientemente la propiedad necesaria, a saber: resistencia eléctrica por unidad de longitud. Si pueden obtenerse valores prácticos de resistencia en los circuitos microelectrónicos es gracias a la precisión en la fabricación que ofrece esta tecnología, que hace posible la realización de tiras de material con un área transversal muy pequeña. A pesar de todo, las resistencias con valores elevados resultan de fabricación engorrosa.



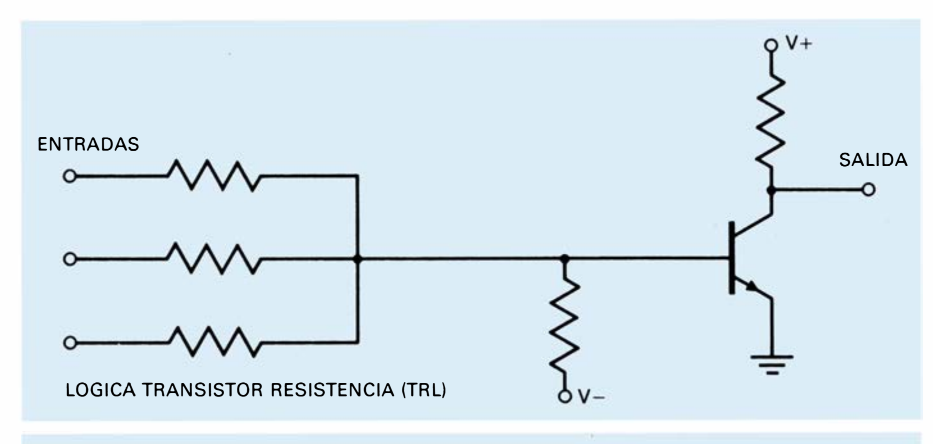
LOS DISPOSITIVOS MOS COMPLEMENTARIOS abarcan en una pastilla de silicio los transistores MOS *n* y los transistores MOS *p*. Si el circuito se fabrica en un substrato de tipo *n*, como el aquí representado, entonces puede construirse un transistor de canal *p*; ahora bien, un dispositivo de canal *n* exige una isla de material de tipo *p*. La necesidad de dicha isla suma otra etapa en el proceso de fabricación y reduce la densidad del encapsulado, y, por consiguiente, el número de transistores que pueden incluirse en una única pastilla. Los dispositivos MOS complementarios pueden ordenarse de suerte que consuman poca energía. En el circuito representado, las puertas de ambos transistores aparecen conectadas a una entrada única; puesto que los dos transistores necesitan señales de polaridad inversa para la conducción, nunca se activan a la vez, con lo cual se consigue un reducido flujo de corriente de la fuente de alimentación (*V*+) a tierra.

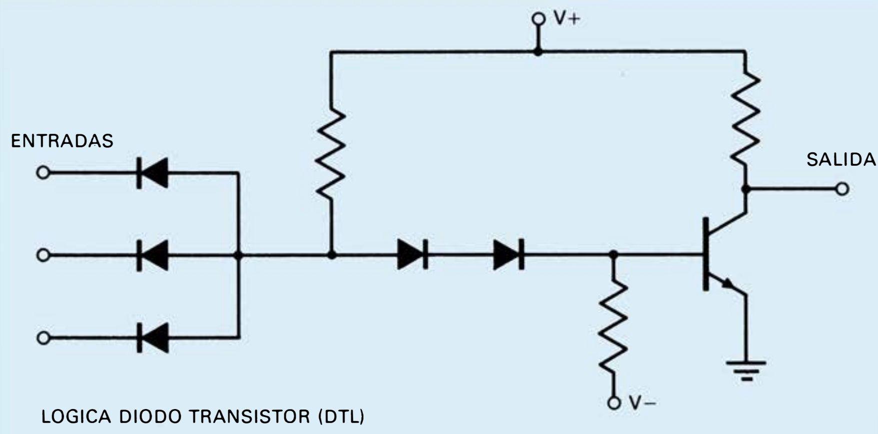
Otro de los inconvenientes de las resistencias de silicio estriba en que el valor definitivo no puede preverse con la precisión requerida; por tanto, los circuitos que incorporen resistencias deberán diseñarse para tolerar variaciones considerables. La fabricación de todo un circuito integrado en una sola serie de procesos puede compensar parte de tan amplias tolerancias. Puesto que todas las resistencias de una pastilla se forman al mismo tiempo, tienden a apartarse de los valores especificados por la misma cantidad aproximadamente; además, los valores de la resistencia varían de consuno con algunas propiedades de los transistores que se incluyen en la misma pastilla. Haciendo que la desviación del valor nominal de un componente se equilibre con el de otro, se obtendrá un rendimiento excelente del circuito en su conjunto.

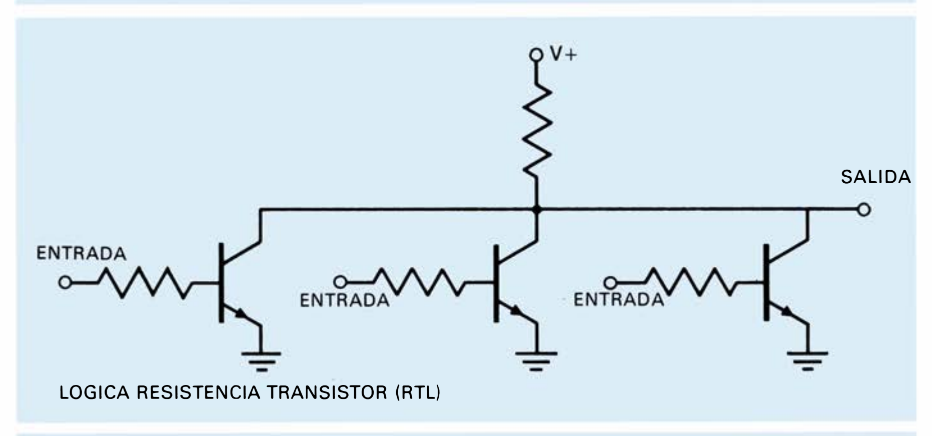
La estructura del condensador microelectrónico es muy semejante a la del electrodo de la puerta de un transistor de óxido metálico semiconductor. Cuando el dispositivo funciona como condensador, sin formar parte de un elemento activo del circuito, se impurifica fuertemente el área de silicio que está situada debajo del electrodo, para aumentar la conductividad. Luego, se añade la capa fina de óxido, a la que sigue otra de aluminio. Al igual que con las resistencias, el método no se emplea en la fabricación de componentes discretos. A fin de evitar cortocircuitos, la capa de óxido de silicio no puede menguarse tanto como podría hacerse con otros aisladores. Además, el área total que presentan los conductores es muy pequeña. Los condensadores discretos están constituidos por numerosas hojas de aluminio intercaladas, en tanto que el equivalente microelectrónico está limitado a un simple par de placas. Por estas razones, los valores de capacitancia disponibles en los circuitos microelectrónicos son pequeños, y rara vez alcanzan valores superiores a la diezmilésima parte del valor de un condensador típico de un circuito que esté montado con componentes discretos.

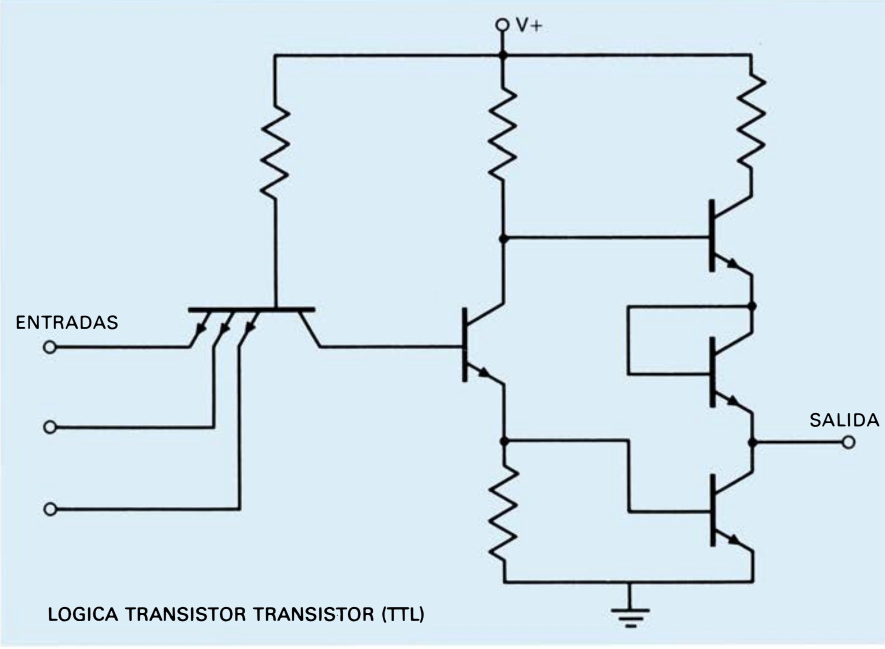
Si cuesta adaptar los condensadores al tamaño de los dispositivos microelectrónicos, los inductores se han mostrado absolutamente reacios a una adaptación. Las bobinas helicoidales y los materiales de núcleo ferromagnético, que forman parte de los inductores discretos, resultan totalmente incompatibles con la tecnología del silicio, sin haberse encontrado aún substituto. Por tanto, cuando se requiera un inductor, habrá que procurarlo como componente discreto.

Los avances en la tecnología aumentarán indudablemente la selección de los elementos pasivos disponibles por parte de los diseñadores de circuitos microelectrónicos. Pero las posibilidades de









mejora aumentan todavía más con la simple eliminación de muchos componentes pasivos en favor de los activos.

El caso de las resistencias resulta particularmente espectacular, ya que con frecuencia se eliminan por substitución directa. El transistor puede contemplarse como una resistencia controlada por corriente o por tensión. De lo cual se deduce que, si se dispone de una corriente o tensión adecuadas para definir la resistencia, el transistor puede conectarse a casi todos los puntos del circuito en donde podría emplearse el dispositivo resistor. Este método jamás podría aplicarse a un circuito montado con componentes discretos, ya que los transistores son más caros que los elementos resistivos discretos. Por otra parte, el coste de un elemento de circuito microelectrónico viene dado por el área de silicio que ocupa; ésta es por lo menos igual para la resistencia que para el transistor.

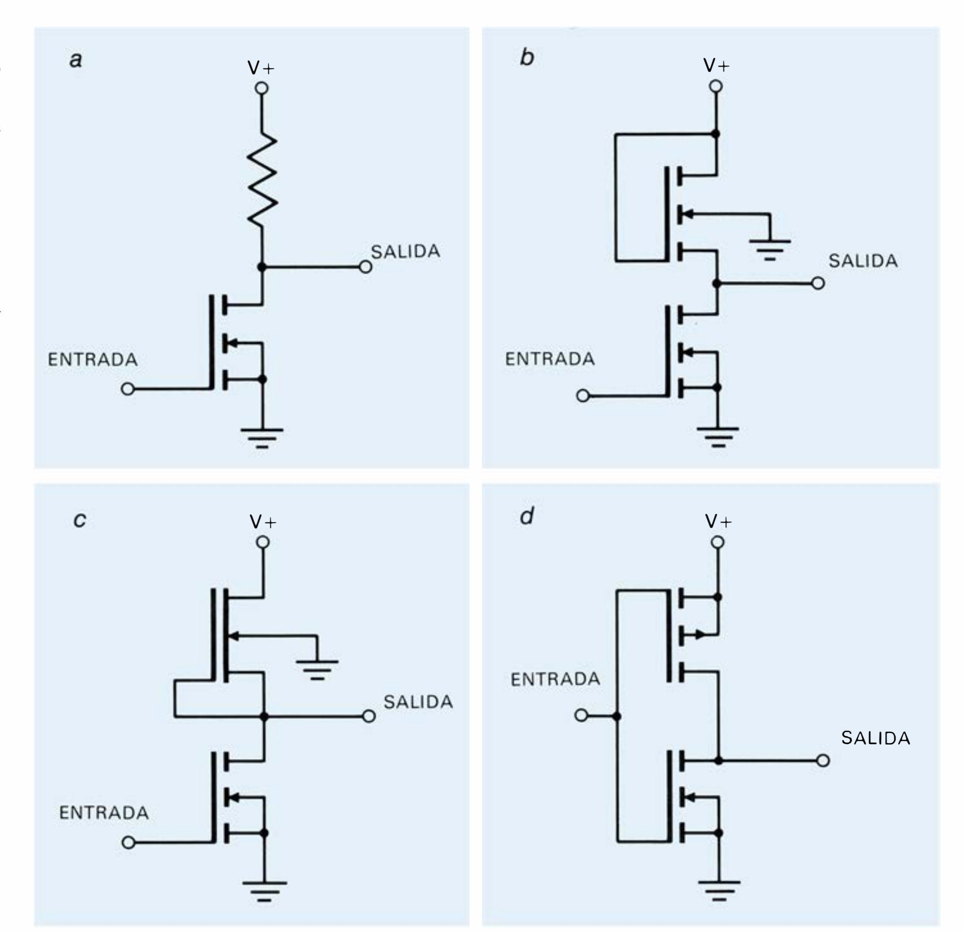
La substitución de otros elementos pasivos del circuito por elementos activos requiere unas técnicas mucho más elaboradas. En cierto sentido, la más elaborada de todas es el súbito predominio de la electrónica digital en los últimos años. Capacitancias e inductancias de valores elevados se emplean en la circuitería convencional para asegurar la exacta reproducción de las tensiones y frecuencias de variación constante de las señales analógicas. Incluso en tales circuitos pueden reemplazarse los grandes elementos pasivos por una red de activos; pero los elementos pasivos se omiten con mayor facilidad en los procesamientos de señales digitales, las cuales presentan sólo niveles discretos. En la mayoría de los

LOS CIRCUITOS LOGICOS digitales que emplean dispositivos semiconductores bipolares han evolucionado hacia un estado en el que los transistores se hallan adaptados para desempeñar casi todas las funciones. La lógica digital opera con señales que sólo tienen dos niveles identificables (verbigracia, alta y baja tensión); los circuitos lógicos aceptan tales señales como entradas y las transforman luego de acuerdo con una normativa fija para generar después la señal de salida. Las primeras familias de circuitos lógicos bipolares se construyeron con componentes discretos. En la lógica transistor-resistencia (TRL), se aumentó al máximo el número de resistencias, puesto que resultaban los elementos más económicos. Por su parte, en la lógica diodo-transistor (DTL) se aumentó el rendimiento al substituir las resistencias por diodos semiconductores. La lógica resistencia-transistor (RTL) fue la primera en emplear tecnología microelectrónica; se suministró un transistor por entrada y se echó mano de sólo unas pocas resistencias con valores bajos. Por último, la lógica transistor-transistor (TTL) es la más común de la tecnología microelectrónica bipolar de nuestros días. Presenta un elevado número de transistores, acoplados directamente entre sí. El circuito representado ofrece un dispositivo, el transistor de emisor múltiple, sin equivalente entre los componentes discretos.

casos existen sólo dos valores reconocibles: alta tensión y baja tensión (o tensión nula). En tales circuitos, los transistores desempeñan el cometido de interruptores, los cuales presentan también dos estados: conexión y desconexión; y, usados de este modo, requieren menos elementos auxiliares.

Los dos tipos fundamentales de transistores, bipolar y MOSFET, dividen a los circuitos microelectrónicos en dos grandes familias. Los dispositivos bipolares se desarrollaron en primer lugar, posibilitando la evolución de una gran variedad de tecnologías bipolares. Hoy en día, la mayor densidad de elementos de circuito por pastilla se ha conseguido con la más reciente tecnología MOSFET, pero ello no significa que los circuitos bipolares estén en trance de quedar obsoletos.

En la construcción de un circuito integrado bipolar típico, se somete primero la pastilla a un proceso de impurificado total, mediante un impurificador de tipo *p*. En este océano de tipo *p* se crean islas de silicio de tipo *n* separadas para cada transistor (y otros elementos del circuito, las resistencias, por ejemplo). Las islas desempeñan la función de aislamiento, ya que el circuito ha sido diseñado para mantenerlas a una tensión positiva respecto al substrato, de suerte que no exista paso alguno de corriente entre ellos, a excepción de una pequeña corriente inversa de diodo.



LOS CIRCUITOS LOGICOS MOS se desarrollaron después de los bipolares, y, en su inmensa mayoría, se construyen en forma microelectrónica. Cada uno de estos circuitos realiza la función de invertir la señal, de suerte que si la entrada es elevada, la salida será baja, y viceversa. En cada caso, la entrada se aplica a la puerta de un transistor MOS *n* incrementado; los circuitos se distinguen por la elección del elemento de carga que se necesita para limitar la corriente a través de dicho transistor. La elección más sencilla es la que ofrece una resistencia *(a)*, pero también es la que registra el peor rendimiento. El segundo transistor MOS *n* incrementado *(b)* es el de fabricación más cómoda; un tercer transistor MOS *n* de agotamiento *(c)* es el que brinda mayor densidad de encapsulado. Por último, si se añade un dispositivo de canal *p* incrementado *(d)*, puede llegar a obtenerse fácilmente el bajo consumo que aporta el circuito MOS complementario.

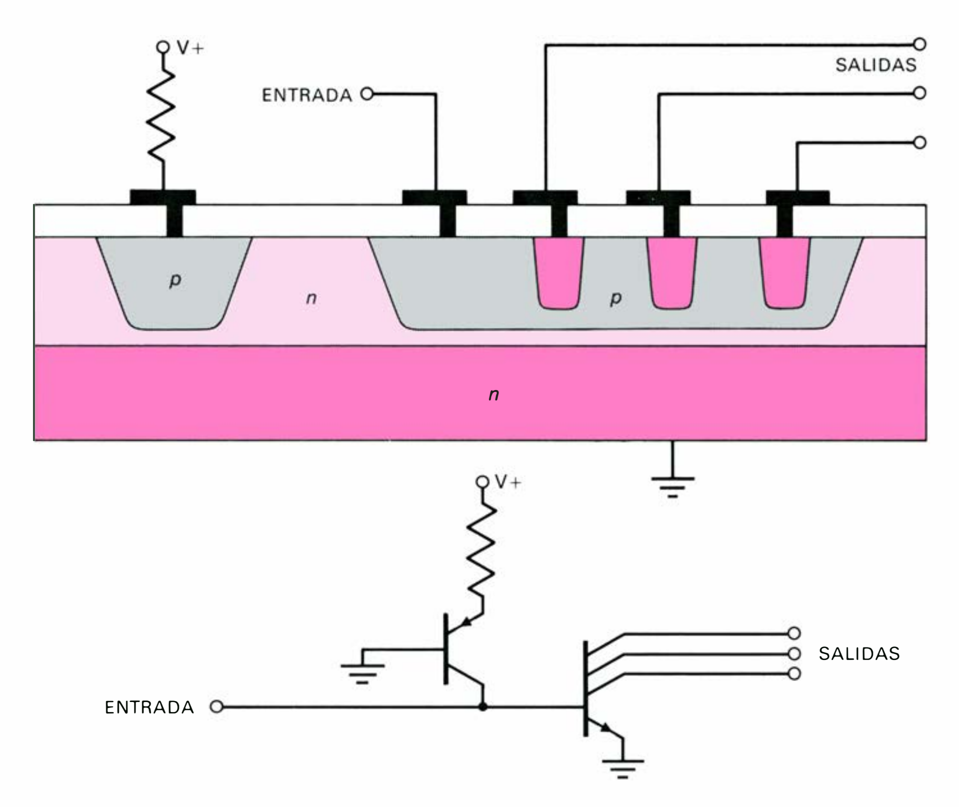
Para confeccionar un transistor *npn*, se introduce un lago de material de tipo *p* en la isla de tipo *n*, creándose finalmente una isla de tamaño menor de semiconductor tipo *n*, que se coloca dentro del lago. La isla más pequeña e interior corresponde al emisor; el lago de tipo *p* es la base y la isla principal, que permanece como aislante, pertenece al colector. Los electrones que forman la corriente del colector descienden desde el emisor, a través de la base, hacia el volumen del colector, relativamente grande.

En una misma pastilla podría construirse un transistor *pnp* de diseño similar al anterior, pero se requeriría de una operación más de impurificado, que vendría a añadirse a las otras cuatro operaciones anteriores. Siguiendo tal proceso, se crearía una isla de material de tipo *n* en el mar contaminado *p*. Luego podría añadirse, sucesivamente, un lago de semiconductor de tipo *p*, para una isla menor de material de tipo *n* y, por último, un pequeño pozo de silicio de tipo *p* en el interior de la isla pequeña. De todas estas regiones, solamente las tres últimas tomarían parte en el funcionamiento del transistor; la isla principal únicamente se emplearía como material aislante.

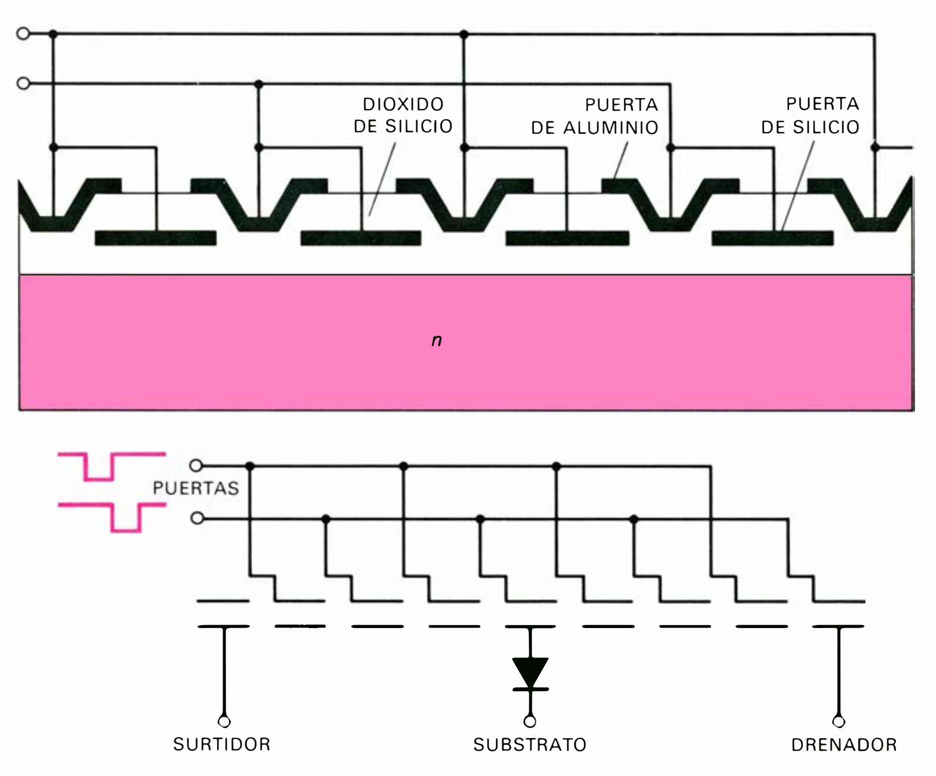
A fin de evitar tales complejidades, el transistor *pnp* se construye de acuerdo con un enfoque distinto, en el cual el flujo de corriente es lateral, y no vertical. En este diseño, la isla aislante de tipo *n* pasa a ser la base, introduciéndose dos regiones *p* separadas. Estas suelen presentar una disposición anular, con el colector formando un anillo en torno al emisor. Las características del transistor lateral *pnp* son inferiores a las del vertical *npn*, porque la región de base que separa al emisor del colector resulta inevitablemente más ancha. A consecuencia de lo cual quedan menguadas la ganancia de corriente y la velocidad de conmutación del transistor lateral.

Pueden también fabricarse transistores MOS en la superficie de un océano de silicio tipo *p*, si bien los circuitos MOS no requieren de diodos de aislamiento. Para construir un MOSFET de canal *n* incrementado basta con crear dos pequeñas islas de material de tipo *n* que sirvan como surtidor y drenador; luego, deben formarse la capa de óxido y el electrodo de la puerta en la superficie de la pastilla y sobre el canal. Las islas aislantes no resultan necesarias, ya que las tensiones que se aplican al surtidor y al drenador aseguran que no haya paso de corriente entre ellos y el substrato. El transistor se autoaísla de los otros elementos de la pastilla.

Un dispositivo microelectrónico de canal *n* destinado a operar en agotamiento está constituido de igual forma que el anterior, con la salvedad de la presencia de un canal permanente de silicio impurificado de tipo *n*, proyectado entre el surtidor y el drenador. Los dispositivos MOS que emplean el canal *p* están formados por un substrato tratado con impurificador de tipo *n*; por tanto, las polaridades de todas las regiones semiconductoras, así como las polari-



LA LOGICA DE INYECCION INTEGRADA, o I^2L, comprende un circuito lógico completo, formado por dos transistores en una sola unidad. El substrato actúa como emisor de un transistor *npn (derecha),* en el cual la corriente fluye ascensionalmente por la base hacia los colectores múltiples. El substrato hace también funciones de base de otro transistor, un dispositivo *pnp* en el que la corriente fluye lateralmente. Merced a esta disposición de los elementos, las islas aislantes que se requerirían normalmente en la tecnología bipolar quedan eliminadas y pueden conseguirse unas densidades de encapsulado parecidas a las que son características de los circuitos MOS.



EL DISPOSITIVO DE ACOPLO DE CARGAS es un elemento del circuito microelectrónico cuya función no podría duplicarse mediante el ensamblaje de componentes discretos. El dispositivo puede considerarse como un transistor MOS exagerado, dotado de una larga hilera de puertas (unas 1000) entre el surtidor y el drenador. En el dispositivo de canal *p* que aquí figura puede retenerse, por breve tiempo, un paquete de carga eléctrica (formado por una concentración de huecos) mediante una tensión negativa constante que se aplica a una de las puertas. Si la tensión remite luego, y se activa simultáneamente la puerta siguiente de la hilera, se registrará un desplazamiento del paquete de carga hacia una nueva posición debajo de la segunda puerta. Aplicando pulsos a puertas alternas se transfiere una secuencia de paquetes de carga del surtidor al drenador.

dades de todas las tensiones aplicadas, deberán estar invertidas.

En los circuitos microelectrónicos que usan tecnología MOS, decidirse por el empleo de dispositivos de canal *n* o *p* resulta, normalmente, fundamental, porque todos los transistores que se incluyen en la misma pastilla utilizan el mismo tipo de material para el canal. No obstante, se pueden realizar ambos tipos de MOSFET en una sola pastilla, a costa de incrementar la complejidad y disminuir su densidad. Los circuitos complementarios MOS se fabrican aislando todos los dispositivos de canal *p* dentro de islas de silicio de tipo *n.* Los circuitos integrados complementarios resultan de fabricación más difícil, pero su mayor coste queda compensado en determinadas aplicaciones. Los dispositivos MOS complementarios (CMOS) presentan la ventaja de un bajo consumo energético.

Al no exigirse islas aislantes en los MOS *n* y MOS *p* convencionales, presentan una densidad de encapsulado mayor que los bipolares. La diferencia es substancial: los transistores MOS ocupan aproximadamente un cuarto de la superficie de silicio que ocuparían los bipolares. Por esta razón, la tecnología MOS impera en la fabricación de circuitos de alta escala de integración, en los que decenas de miles de elementos activos pueden incorporarse en un solo dispositivo. Por ello, al igual que casi todas las memorias de semiconductor, la mayoría de microprocesadores de una pastilla se construyen en tecnología MOS.

Una de las prestaciones más importantes que presentan los circuitos bipolares es su elevada velocidad de operación. La velocidad a la que un transistor bipolar puede cambiar su estado está determinada principalmente por la amplitud de la base. El tiempo de retardo viene expresado por el tiempo que emplean los electrones en difundirse del emisor al colector. En los dispositivos de diseño vertical, utilizados generalmente para transistores *npn,* la base puede ser bastante estrecha, mucho más que la línea más estrecha que pueda determinarse fotolitográficamente en la superficie de la pastilla. La amplitud de la puerta en los dispositivos MOS desempeña un papel similar; ahora bien, puesto que los transistores MOS son dispositivos laterales, el ancho no puede reducirse por debajo del límite de resolución. La movilidad de los portadores de carga afecta también a la velocidad del transistor MOS, ya que los electrones se mueven con mayor rapidez que los huecos; por tanto, los dispositivos de canal *n*

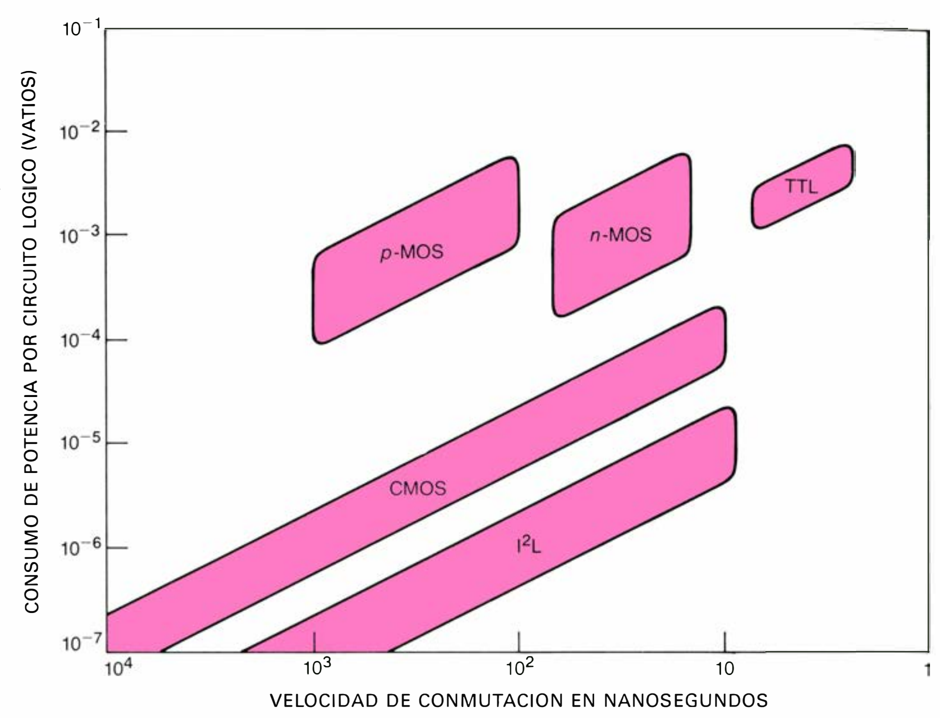
tienden a ser más veloces que sus homónimos de canal *p*. Finalmente, cabe decir que la velocidad de los dispositivos MOS viene condicionada por el condensador formado por la puerta, por la capa de óxido y por el substrato, que necesita un tiempo limitado para la carga y descarga. La reducción de capacitancia acelera el funcionamiento pero recorta también la ganancia.

El desarrollo de los circuitos microelectrónicos en las dos últimas décadas ha hecho que hubiese transistores de elevada calidad a precios cada vez más baratos y en mayor abundancia, mientras que sólo ofrecía una pobre selección de resistencias y condensadores y absolutamente nada en lo referente a inductores. De lo cual resultó que las configuraciones de los circuitos evolucionaron en el sentido de aprovechar las ventajas de la tecnología y evitar sus limitaciones. Puede seguirse la tendencia a través de la evolución de los circuitos lógicos digitales.

Las unidades fundamentales de los circuitos lógicos se denominan puertas (que no deben confundirse con los electrodos de igual denominación existentes en los transistores MOS). Cada tipo de puerta acepta en sus terminales de entrada una señal y la transforma de acuerdo con una regla incorporada en su cableado interno; la señal transformada aparece luego en los terminales de salida. Hay puertas que aceptan más de una entrada y determinan la señal de salida con arreglo a la combinación de las señales de entrada. El corazón de cada puerta está formado por un elemento activo al menos, que actúa como conmutador: este conmutador controla la salida del circuito y, a su vez, es controlado por la señal de entrada.

Entre las familias de circuitos lógicos desarrollados a mediados de los años 50, figura la lógica transistor-resistencia, TRL (del inglés "transistor-resistor logic"). Los circuitos se montaban con componentes discretos aumentando al máximo el número de resistencias debido a su bajo precio y gran fiabilidad; un ejemplo de puerta típica de tres entradas consistía en el montaje de un transistor y cinco resistencias.

Más tarde, a finales de los años 50 y a comienzos de la década de los 60, los diodos semiconductores, encapsulados también como componentes discretos, adquirieron un precio que les permitía competir con las resistencias; resultado de ello sería el nacimiento de la lógica diodo-transistor o DTL (del inglés "diode-transistor logic"), en la que los diodos se empleaban para aislar las señales de entrada así como para desplazar los niveles de tensión. Una puerta típica estaba formada por un transistor conmutador, tres resistencias y cinco diodos. Tales montajes presentaban un bajo consumo de potencia y velocidad más elevada que la ofrecida por la red equivalente transistor-resistencia.



LA ELEVADA VELOCIDAD Y EL BAJO CONSUMO en los dispositivos microelectrónicos presentan una relación recíproca, ya que el circuito puede hacerse más rápido con sólo aplicarle mayor potencia. El producto del tiempo de retraso de conmutación por el consumo resulta de gran importancia en los circuitos digitales, pues representa la energía requerida en una operación de conmutación. Actualmente los circuitos lógicos más rápidos emplean tecnología bipolar, al igual que los TTL, pero presentan un mayor consumo de energía. En tecnología MOS, los dispositivos MOS *p* están siendo reemplazados por los MOS *n*, porque estos últimos operan con mayor rapidez sin necesidad de mayor consumo. Los circuitos MOS complementarios ofrecen menor consumo sin detrimento de la velocidad. Finalmente, la tecnología I^2L resulta de gran interés si tenemos en consideración que presenta gran velocidad y bajo consumo.

La siguiente generación de diseños lógicos, generalizada a principios y mediados de los años sesenta, introdujo los primeros circuitos lógicos microelectrónicos. A esta familia se la conoció como lógica resistencia-transistor (o RTL, del inglés "resistor-transistor logic"); las resistencias tenían valores muy inferiores a los de las empleadas en los circuitos primitivos y los transistores se utilizaban con mucha más generosidad. En una puerta típica de este tipo, cada una de las tres entradas tenía asignado su propio transistor interruptor. Volvieron a obtenerse avances en el consumo y en la velocidad. Los circuitos integrados RTL estaban fabricados con una o varias puertas en una pastilla.

La cuarta familia lógica que había de sumarse permanece todavía operante; se trata de la más común de toda la tecnología digital bipolar. Se la reconoce por lógica transistor-transistor (o TTL, del inglés "transistor-transistor logic"), y, como se deduce del nombre, delega la mayoría de funciones en los elementos activos del circuito. Las puertas TTL fueron las primeras con importancia comercial que incorporaron un elemento de circuito microelectrónico que no hubiese podido montarse con componentes discretos. Dicho elemento es el transistor de emisor múltiple, en el cual dos o tres emisores comparten una base y colector comunes. En una puerta TTL, cada uno de los emisores acepta una señal de entrada. El transistor de emisor múltiple controla un transistor conmutador único, que, a su vez, excita una red de tres transistores de salida. Comparados con la familia anterior de puertas RTL, las circuitos TTL presentan mayor potencia de salida (de suerte que puede excitarse un mayor número de puertas en la etapa siguiente), mayor tolerancia de fabricación y mayor inmunidad a las tensiones espúrias, o ruido eléctrico. En la actualidad, resulta habitual encontrar circuitos integrados que contengan varios cientos de puertas TTL.

Si la evolución de la tecnología MOS ha sido menos complicada, quizá se deba a que es más tardía. Por lo general, las familias lógicas modernas no han

desplazado por completo a las antiguas, ya que cada una de ellas se adapta mejor a un determinado tipo de aplicaciones.

Los primeros circuitos microelectrónicos MOSFET empleaban exclusivamente elementos de canal *p*, debido a que presentaban mayor sencillez de fabricación. Pero también son más lentos que los otros tipos, a causa de la menor movilidad de los huecos respecto a los electrones. Los dispositivos MOS de tipo *p* han sido desplazados en las aplicaciones que requieren un rendimiento elevado; no obstante, todavía continúan empleándose mucho en los circuitos de las calculadoras electrónicas de bajo coste, en las cuales el precio resulta más importante que la rapidez de operación.

Los circuitos MOS de canal *n* han reemplazado a sus homónimos de canal *p* en los microprocesadores y en memorias de semiconductor. Las variedades de la tecnología MOS de canal *n* podemos examinarlas si consideramos un simple circuito de puerta con una entrada a un transistor de efecto de campo incrementado de canal *n*. En él, las posibles configuraciones del circuito vendrán determinadas por el elemento elegido como carga a adaptar a este transistor.

La elección más directa y conceptualmente más simple es la de una resistencia, aunque también es el elemento que presenta peores rendimientos. Otra posibilidad para obtener la carga limitadora de la corriente es el uso de un transistor de efecto de campo incrementado de canal *n*, de tipo similar, conectado en serie con el primero. Con ello se obtiene un proceso de fabricación más sencillo y es la combinación de empleo más generalizado en aplicaciones de bajo coste. La mayor densidad de encapsulado se consigue en los dispositivos de agotamiento de canal *n*, que actúan como carga; por ello se usan como circuitos de alta escala de integración en aplicaciones de microprocesadores. Por último, puede utilizarse un transistor incrementado de canal *p* como elemento de carga; la puerta deja de ser un dispositivo MOS *n* para convertirse en un circuito complementario MOS (CMOS). Porque sólo uno de los transistores conduce en un tiempo dado, excepto en el momento de la conmutación, hay poca corriente estacionaria a través de la puerta, por lo que el consumo será bajo. La principal aplicación de la tecnología CMOS se encuentra en la construcción de los relojes de pulsera electrónicos, que funcionan mediante diminutas baterías durante períodos de tiempos muy largos.

Merced a los avances en los métodos de fabricación y en la ciencia de los materiales, la densidad de encapsulado de los dispositivos bipolares ha venido registrando notables mejoras, e igual ha pasado con la velocidad de los dispositivos MOS. Sin embargo, existe una laguna substancial entre los mismos. Para salvarla se ha creado una nueva tecnología denominada lógica integrada de inyección, normalmente conocida por lógica I²L (del inglés "integrated-injection logic").

El elemento funcional básico de la tecnología I²L consta de dos transistores que forman una unidad en el substrato de silicio. Uno de ellos presenta una disposición vertical, pero la corriente asciende hacia el colector, al revés de como sucede en los diseños de uso corriente. El substrato hace de emisor de este transistor, y, al mismo tiempo, colabora en la formación de la base del otro transistor lateral. A causa de las tensiones que se aplican a las varias regiones impurificadas, este tipo de dispositivos no necesita islas aislantes; y la pareja constituida por estos transistores "superintegrados" forma una propia puerta lógica completa.

La velocidad de las puertas de tecnología I²L no alcanzará probablemente a la de los circuitos TTL más rápidos, pese a que la I²L es ya más rápida que todas las tecnologías MOSFET. Por otra parte, la arquitectura compacta que presenta la lógica I²L la convierte en una de las más firmes candidatas para la integración a gran escala. Además, resulta viable que las densidades de encapsulado sean iguales a las de las pastillas MOS de canal *n*.

Los dispositivos que hemos considerado forman parte de la corriente principal del desarrollo de la microelectrónica. Sin embargo, las técnicas de fabricación de las puertas lógicas se han aplicado al desarrollo de otros tipos de dispositivo electrónico. Los circuitos que procesan señales analógicas son probablemente indispensables en ciertos dispositivos, teléfono y radio por ejemplo, pudiendo alguno de ellos fabricarse en forma microelectrónica. Un amplificador típico de baja potencia diseñado para el montaje con componentes discretos tiene un solo transistor, si bien exige cuatro resistencias, dos de las cuales presentan valores elevados, y tres condensadores cuyos valores se hallan entre las gamas más elevadas. Un dispositivo microelectrónico equivalente emplea siete transistores con una sola resistencia controladora de corriente. En determinados

casos, un sistema más eficaz para tratar señales analógicas puede ser convertirlas en digitales, procesarlas y reconvertirlas luego en analógicas, cuando sea necesario. El aparato para la conversión puede fabricarse también en un encapsulado microelectrónico.

La tecnología de los semiconductores ofrece también una selección de transductores a través de los cuales los dispositivos electrónicos pueden comunicar con el entorno. Un ejemplo reciente de ello son los medidores de presión, cuyo elemento activo está constituido por un fino diafragma de silicio. La resistencia del silicio varía en respuesta a la presión mecánica; de esta forma, los cambios de presión a través del diafragma se detectan como variaciones en la resistencia del material.

Los transductores que resultan algo más familiares son los fotodiodos y los fototransistores. Cuando un fotón, o cuanto de energía electromagnética, es absorbido por la región de tipo *p* de un fotodiodo, la energía cedida por el fotón crea un electrón y un hueco. El electrón se dirige hacia la unión, cruzándola; la corriente total de esos electrones es proporcional al flujo de fotones incidentes.

El circuito inverso al anterior es el diodo luminiscente, que está diseñado para aprovechar las colisiones entre electrones y huecos. En cada colisión, el electrón llena un hueco, y ambas partículas se aniquilan en cuanto portadoras de carga. Su energía aparece en forma de un fotón. En el silicio, las aniquilaciones conducen a la emisión de radiación infrarroja; sin embargo, en la mezcla semiconductora de arseniuro de galio, la luz irradiada se encuentra en la zona roja del espectro visible. Conjuntos de diodos emisores de luz de arseniuro de galio forman parte de los visualizadores numéricos de muchas calculadoras.

Existe un elemento microelectrónico funcional de reciente invención que presenta la particularidad de que ningún montaje de componentes discretos puede emular su funcionamiento; a estos depuradísimos elementos microelectrónicos se les denomina dispositivos de acoplamiento de carga. Pueden considerarse como una extensión del transistor de efecto de campo MOS incrementado. El dispositivo de acoplamiento de carga consta de un surtidor y un drenador, separados por una larga hilera de puertas.

Un impulso de carga inyectado por el surtidor puede atraparse en la primera puerta gracias a la tensión aplicada a este electrodo. Si la tensión de la primera puerta baja, y se aplica simultáneamente otra tensión a la segunda puerta, tendremos que la carga es atraída y atrapada en la segunda puerta. Y así, repitiendo el proceso, se transfiere la carga de puerta en puerta hasta que alcanza al drenador.

Pueden fabricarse memorias en las que los dígitos se almacenen en forma de largas secuencias de bits construyendo dispositivos de acoplamiento de carga de unas 1000 puertas. Puesto que la densidad de encapsulado puede llegar a ser muy alta, pueden incluirse gran cantidad de tales dispositivos en una sola pastilla. Los impulsos de carga pueden generarse fotoeléctricamente en vez de inyectarlos por el surtidor; de suerte que un conjunto de dispositivos de acoplamiento de carga puede funcionar como el corazón de una cámara de televisión con salida digital.

Quedó atrás el tiempo en que el diseñador de un circuito microelectrónico podía estudiar el comportamiento de su ejemplar por medio de una maqueta prototípica realizada a partir de componentes discretos. La forma directa de diseño consiste en representar matemáticamente el circuito, estudiando sus características con ayuda de computadores.

A su vez, el modelo debe basarse también en una detallada descripción de las distintas propiedades eléctricas de cada elemento del circuito. A medida que las dimensiones de los elementos se apartan cada vez más de la escala humana, se hace cada vez más difícil la tarea de verificar tales descripciones. Al ser las dimensiones de los elementos inferiores a una micra, efectos físicos indiferentes para dispositivos mayores pueden desempeñar un papel de gran importancia, a veces fundamental. Por consiguiente, un requisito para el progreso incesante que no hay que dejar de lado es la necesidad de mejorar los modelos de los elementos del circuito.

En la fabricación de la próxima generación de circuitos integrados será necesaria también una aproximación sistemática. Los modelos no sólo habrán de revelar cómo operan los elementos, sino también cómo hay que fabricarlos. Si se comparan con los modelos de comportamiento eléctrico, estos modelos de proceso se encuentran en un estado primitivo; la fabricación sigue siendo como un arte empírico. El avance continuo en el ámbito de la microelectrónica puede depender grandemente de la capacidad de predecir las propiedades de un transistor a partir del conocimiento de los pasos seguidos en su fabricación.

Integración a gran escala en los circuitos microelectrónicos

Se dice que la integración es a gran escala cuando miles de circuitos electrónicos básicos están integrados en una pastilla. Muchos de estos circuitos, además de ser eléctricos, siguen las reglas de la lógica booleana

William C. Holton

La velocidad a la que se pueden realizar los cálculos ha aumentado en los últimos 20 años en seis órdenes de magnitud, es decir, en un millón de veces, aproximadamente. Y lo que es más, el coste de la computación ha bajado drásticamente. Minicomputadores que hoy en día se pueden adquirir por 1000 dólares vienen a igualar las capacidades que poseían las grandes máquinas cuyo coste era de 20 millones de dólares hace 15 años. Hacia 1985, un minicomputador de escala media valdrá unos 100 dólares. Por otra parte, el consiguiente aumento de nuestras posibilidades de procesar información nos lleva a una revolución intelectual.

En los años cuarenta, se reconoció la necesidad de aumentar la velocidad y capacidad de cálculo, pero los logros estaban limitados por la tecnología electrónica disponible, vinculada a los tubos de vacío. El calor generado por los tubos de vacío hacía que éstos tuviesen una corta vida operativa, lo cual determinaba una cota superior en el tamaño de los computadores. El primer computador completamente electrónico, ENIAC, se desarrolló en la Universidad de Pennsylvania en 1945; tenía 18.000 tubos de vacío. Otras máquinas que fueran mayores resultaban naturalmente poco prácticas; hubiese sido necesario al menos 24 horas cada día para hallar y reemplazar los tubos defectuosos.

El computador se salvó de éste su fin prematuro gracias a la invención del transistor, en 1947. La reducción de tamaño y coste experimentada en los últimos 10 años se debe al desarrollo del circuito integrado a finales de la década de los 50. El diseño de computadores y otros aparatos electrónicos se está transformando hoy en día de nuevo gracias a la integración a gran escala; mediante este proceso, decenas de miles de transistores y sus interconexiones se fabrican simultáneamente. Merced a esta tecnología, casi todos los elementos lógicos de un computador digital pueden ser encajados en una pastilla de silicio, de no más de 6 mm de lado.

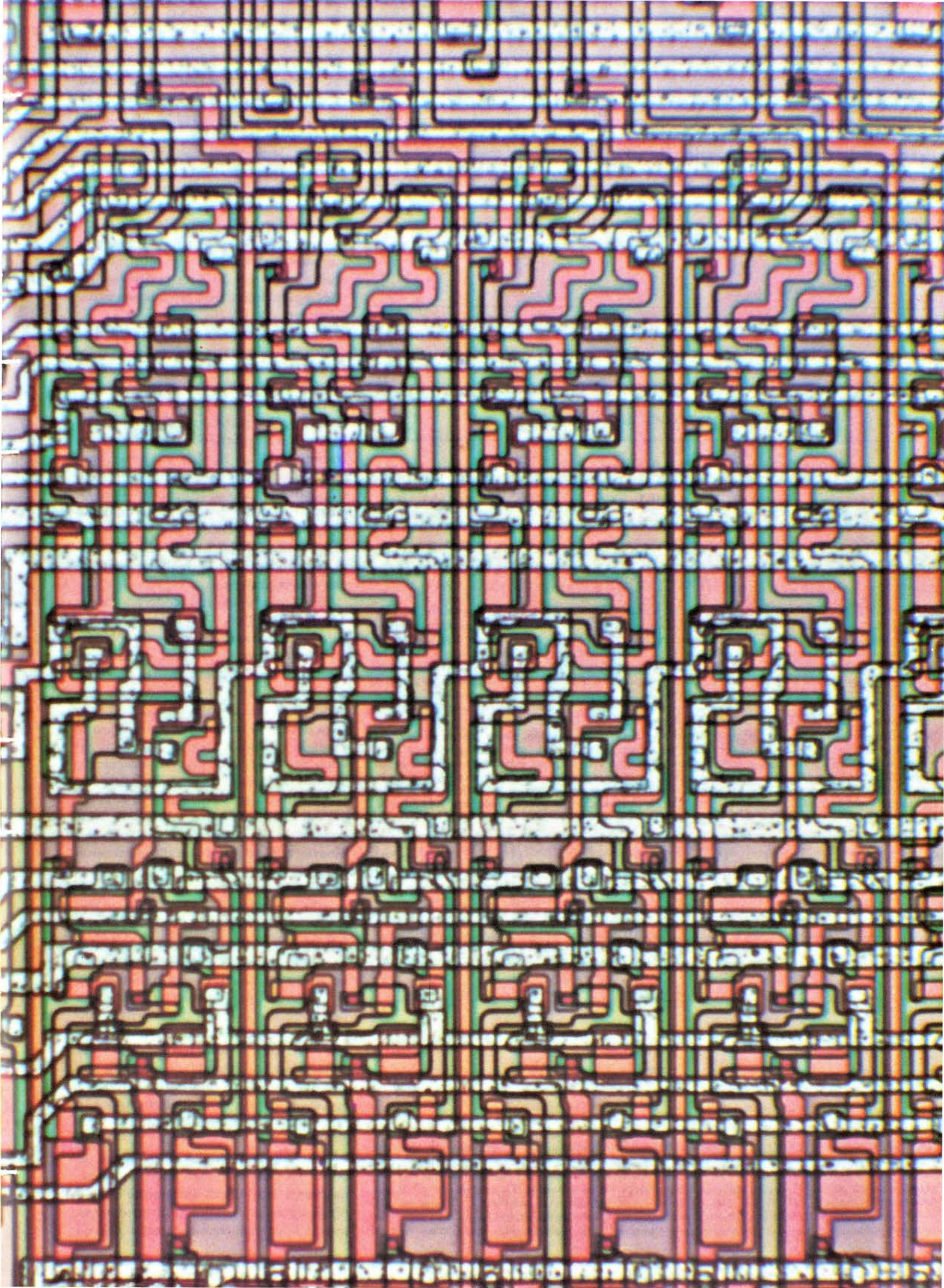
Los transistores de un circuito digital operan a modo de conmutador, que suele presentar solamente dos estados: apagado o encendido, conductor o no conductor. De modo análogo, las señales empleadas en los circuitos tienen únicamente dos niveles de tensión reconocibles, calificados simplemente de alto o bajo. Por hallarse todos los estados de un circuito digital confinados en estos dos valores, las funciones lógicas y aritméticas de un computador se expresan en números binarios, es decir, en números escritos en base 2.

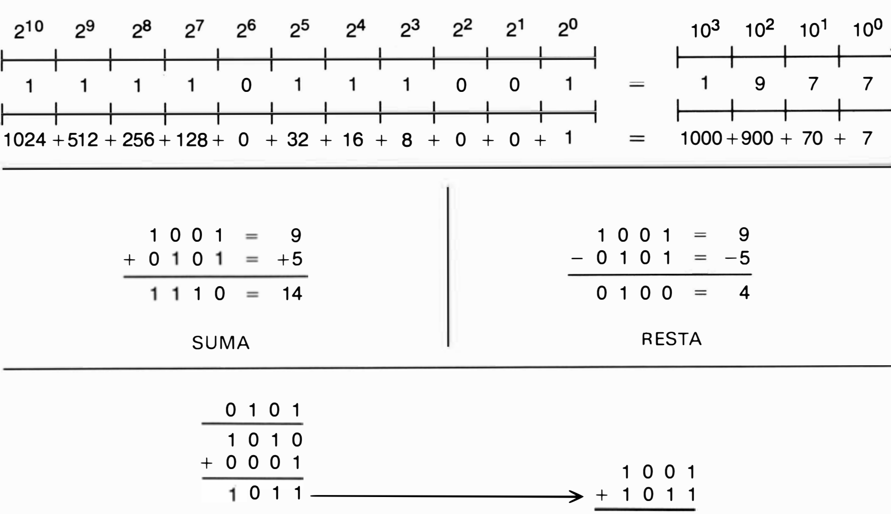
Cuando se lee un número decimal de derecha a izquierda, se entiende que ha de multiplicarse cada dígito por las sucesivas potencias de 10; con frecuencia, solemos denominar a los dígitos como unidades, decenas y centenas, y así sucesivamente. Los números binarios emplean una denominación posicional similar, si exceptuamos que los multiplicandos son potencias de 2; así pues, los dígitos representan unos, doses, cuatros, ochos, y así sucesivamente. Con este sistema de contar, puede representarse cualquier cantidad a partir de una sucesión de ceros y unos.

En el sistema binario, la aritmética resulta en muchos aspectos más sencilla que las correspondientes operaciones con números decimales. La adición consiste en poco más que en el hecho de contar, siempre que se pueda transportar un dígito binario, o bit, hacia adelante, hasta la próxima potencia de dos. La multiplicación es también un procedimiento simple. Dado que cada dígito del multiplicador ha de ser 0 o 1, cada producto parcial formado deberá ser igual a cero o igual al multiplicando. Así pues, la regla para la multiplicación dice que hay que reescribir el multiplicando, desplazado un puesto hacia la izquierda por cada bit igual a 1 en el multiplicador, y sumar luego tales números.

La resta puede hacerse siguiendo el método usual, incluyendo la toma en consideración de bits de la columna inmediatamente superior; aunque también puede realizarse por otra vía que admite una mecanización más adecuada. Esto es, "invirtiendo" todos los bits de un número binario; o lo que es lo mismo, cambiando todos los ceros por unos y todos los unos por ceros, se obtiene un número llamado complementario; éste posee ciertas características de los números negativos. Por consiguiente, se pueden restar números convirtiendo el substraendo en su com-

SUMADOR ELECTRONICO, en la microfotografía de la página opuesta; se trata de una porción pequeña de un microprocesador, íntegramente formado en la superficie de una pastilla de silicio. El microprocesador, que ejemplifica la integración a gran escala, es el TMS 9900, fabricado por Texas Instruments Incorporated. La organización del sumador puede apreciarse en la pauta repetitiva, que divide el área mostrada en cuatro columnas verticales. Cada columna es un circuito, confeccionado con 31 transistores, que operan sobre un dígito, o bit, de los números binarios que han de sumarse. En total hay 16 columnas de éstas; por tanto, el microprocesador puede sumar números de 16 bits. En todos los cálculos aritméticos se emplea el sumador. Las superficies verdes de la pastilla son, principalmente, las regiones del surtidor y el drenador de transistores de óxido metálico semiconductor; las puertas de los transistores aparecen en rosa. Las zonas teñidas en malva reflejan superficies revestidas de dióxido de silicio, y las líneas plateadas son conductores de aluminio.






```
      1 0 0 1   =   9
    × 0 1 0 1   =  ×5
    ---------
      1 0 0 1
    0 0 0 0
  1 0 0 1
0 0 0 0
-------------
0 1 0 1 1 0 1   =  45
```

MULTIPLICACION

LOS NUMEROS BINARIOS pueden representar cualquier cantidad en forma de una sucesión de ceros y de unos. Leídos de derecha a izquierda, ha de entenderse que los números binarios están multiplicados por las sucesivas potencias de dos al igual que los dígitos de los números decimales lo han de ser por las potencias de 10. La suma binaria consiste en poco más que el hecho de contar, con la prevención de transportar un bit hacia adelante, hasta la próxima potencia de dos. Por lo que respecta a la resta, puede seguirse de forma usual, pudiendo sustraerse del bit de la potencia próxima superior; pero puede hacerse también por medio de un número denominado el complementario, que se forma cambiando todos los ceros por unos y los unos por ceros, sumándole a continuación un 1 al resultado. La diferencia entre dos números se halla sumando el complementario del sustraendo al minuendo. La multiplicación está muy simplificada, dado que los dígitos del multiplicador son siempre ceros o unos; por consiguiente, cada producto parcial es igual a cero o igual al multiplicando. Los productos parciales se escriben debajo, desplazados un número conveniente de puestos hacia la izquierda, y luego se suman. La división (no figura aquí), puede realizarse a través de restas repetidas. Todas las operaciones siguen unos procedimientos que la máquina puede realizar sin ninguna dificultad. Además, éstas pueden reducirse a la adición.

A	$\bar{A}$
0	1
1	0

A	B	A · B
0	0	0
0	1	0
1	0	0
1	1	1

A	B	A + B
0	0	0
0	1	1
1	0	1
1	1	1

A → $\bar{A}$ NO

A, B → A · B Y

A, B → A + B O

LAS FUNCIONES LOGICAS también pueden ser expresadas en términos de números binarios. La función "no" "invierte" un dígito binario, cambiando el cero por el uno y el uno por el cero. Las funciones "y" y "o" aceptan dos bits como entradas (designadas por *A* y *B*); generan un bit de salida determinado por los valores de las entradas. La función *A* "y" *B* (escrita simbólicamente *A·B*) es un 1, si y sólo si *A* y *B* son unos. La función "o" (que se escribe *A + B*) genera un uno de salida si o *A* o *B* son uno, o si las entradas de ambas lo son. Las representaciones simbólicas inferiores corresponden a "puertas" que realizan tales operaciones lógicas.

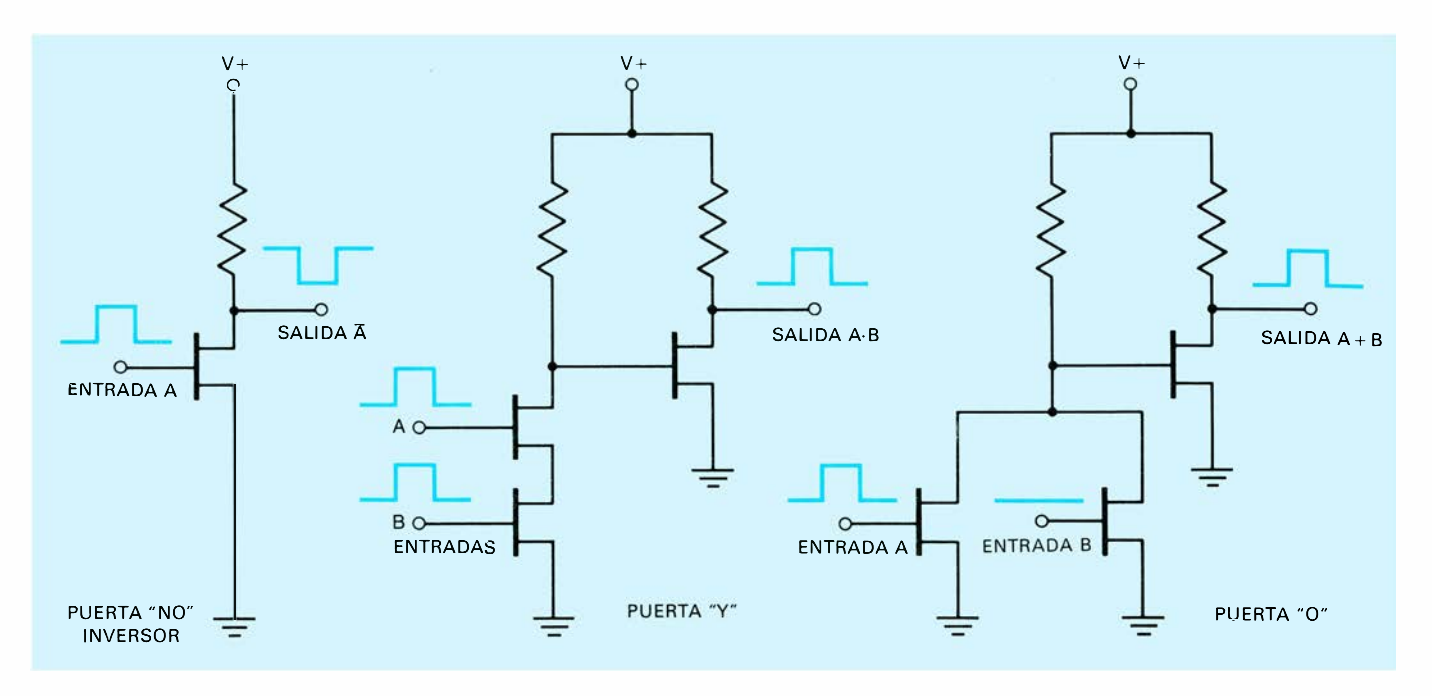
plementario, y siguiendo, a partir de entonces, el esquema propuesto para la adición. Finalmente, para dividir, se cuenta cuántas veces puede substraerse un número de otro, siendo dicho resultado el cociente. Una importante conclusión que se puede extraer de estos procedimientos es que todas las operaciones aritméticas con números binarios pueden reducirse a la adición. Se desprende, por tanto, que un ordenador que no disponga de circuitos especiales para la substracción, multiplicación y división podrá realizar tales operaciones siempre que sea capaz de sumar.

Más importante que la oportunidad de la aritmética binaria sea, tal vez, la facilidad con que los números binarios pueden expresar proposiciones en la lógica simbólica. En 1938, Claude E. Shannon, del Instituto de Tecnología de Massachusetts, destacó que conmutando circuitos, de un tipo de construcción a base de relés electrónicos, éstos podían emplearse para evaluar proposiciones lógicas. En otras palabras, la dualidad existente entre encendido y apagado, tensión alta o baja, 1 o 0, podía adecuarse para representar la dualidad existente entre verdadero y falso.

El sistema lógico empleado por los ordenadores digitales fue ideado por el matemático inglés George Boole un siglo antes, por lo menos, de la construcción del ordenador. El álgebra de Boole ofrece la posibilidad de disponer de un procedimiento riguroso para decidir si una proposición lógica es verdadera o falsa, siempre y cuando dichas proposiciones sean expresables a través de variables que tengan exactamente dos valores posibles. Por convención arbitraria, representaremos lo "verdadero" de la lógica booleana por el 1 binario o bien por el estado de tensión alta.

A través del álgebra de Boole, el análisis lógico se puede ejercer con sólo tres funciones, llamadas "no", "y" y "o". La función "no" cambia un bit binario por su valor opuesto, esto es: convierte un 0 en un 1 y un 1 en un 0. Las funciones "y" y "o" determinan un solo bit de salida a partir de dos o más valores correspondientes a bits de entrada. La función "y" toma el valor "verdadero" únicamente cuando todos los bits de entrada son "verdaderos". Por otra parte, la función "o" toma el valor "verdadero" si por lo menos uno de los bits de entrada es "verdadero".

Las representaciones electrónicas de estas funciones se llaman puertas lógicas. La puerta "no" acepta únicamente un solo bit de entrada y cambia o invierte su valor. Si la entrada es una tensión alta, en



LAS PUERTAS LOGICAS ELECTRONICAS valoran expresiones aritméticas y lógicas en las cuales los valores binarios están representados por tensiones. Por convención un 1 binario está representado por una tensión alta y un 0 por una tensión baja. Las puertas que se ven aquí están construidas con transistores de óxido metálico semiconductor. La más simple es la puerta "no", o inversor. Cuando la entrada a estas puertas está en una fase baja, el transistor no conduce y sólo fluye una corriente despreciable desde la tensión de alimentación (*V*+), a través de la resistencia y del transistor, hasta la masa. Como resultado, hay poca caída de tensión a través de la resistencia; la salida equivale a la tensión de alimentación. Cuando se aplica una señal alta a la entrada, el transistor conduce; la corriente relativamente intensa que circula por el circuito produce una caída considerable a través de la resistencia. El voltaje de salida se asemeja al de tierra y está, en consecuencia, en el estado bajo. La puerta "y" tiene dos transistores de entrada conectados en serie; sólo fluye corriente a través de ellos cuando ambos reciben simultáneamente una señal alta. Para restablecer la adecuada polaridad de la señal, la salida de los dos transistores en serie es seguida por un inversor. En una puerta "o" los transistores de entrada están conectados en paralelo; por tanto, una señal alta que se aplique a uno de ellos da como resultado la conducción. Se requiere otro inversor para cambiar la polaridad de la salida.

la salida es baja; si cambia la entrada, también lo hace la salida. La puerta "y" acepta múltiples señales de entrada; produce una tensión alta de salida sólo si en todas las entradas la tensión es alta. La salida de una puerta "o" es alta siempre y cuando aparezca una tensión alta por lo menos en una de sus entradas.

Las puertas lógicas son unidades funcionales básicas para las operaciones aritméticas y para las operaciones lógicas. Un ejercicio que demuestra efectivamente cómo pueden combinarse las puertas es la construcción de un circuito que sume números binarios. El circuito debe aceptar como entrada a los dos bits que han de ser sumados junto con el bit de transportador, que puede ser 0 o 1, a partir de la potencia de 2 inmediatamente inferior. Debe producir como salida el bit suma y el bit transportador hacia la potencia de 2 inmediatamente superior.

La construcción de la tabla de verdad, que nos da la salida deseada para cada posible combinación de entradas, constituye la primera etapa del diseño del circuito. Hay ocho combinaciones posibles de las tres entradas binarias. El procedimiento para generar el bit suma consiste en hallar todas las combinaciones que requieran un 1 como salida; luego, se ordenan las puertas de forma que se produzca una salida tal. Si los tres bits de entrada son ceros, entonces el bit suma será evidentemente igual a cero. Si únicamente uno de los bits de entrada es igual a 1, y los dos restantes igual a cero, entonces se requiere el 1 para la salida. Existen tres combinaciones de este tipo (ambos sumandos o el bit "transportador" pueden ser unos) y se ha de construir el circuito condicionándolo a cada una de las posibilidades. Si dos entradas son unos, y cero la tercera, entonces el bit suma es cero, dado que en aritmética binaria 1 más 1 es igual a 10. (El bit transportador será generado en un grupo aislado de puertas). Finalmente, los tres bits de entrada pueden ser unos, en cuyo caso la salida será de nuevo un 1 (y de nuevo ha de ser generado un bit transportador). Resumiendo, pues, cuatro de las ocho combinaciones de entradas requieren un 1 de salida; tres combinaciones donde únicamente una de las entradas se halla en el estado 1 y una combinación en la que las tres entradas son unos.

Podemos construir, mediante segmentos, un circuito para el cálculo del bit suma, cada uno de los cuales representaría una de las cuatro combinaciones de "verdadero" de las entradas. Se puede empezar con la combinación en la que los dos sumandos sean ceros y el bit transportador de la columna previa un uno. Esta combinación tiene que dar lugar a un 1 de salida; por tanto, ambos sumandos han de ser previamente pasados a través de una puerta "no", convirtiéndolos en unos; luego se aplican a las entradas de una puerta "y". La salida de esta puerta y el bit transportador (en su forma no invertida) sirven como entradas de otra puerta "y", la cual da el resultado justo.

Se requieren otros dos conjuntos idénticos de puertas. En uno, el bit del primer sumando y el del transportador son invertidos por las puertas "no" y luego combinados con el segundo sumando: en el otro, los bits del segundo sumando y del transportador se invierten y combinan con el primer sumando. Por último, se necesita una cuarta disposición de puertas para producir la salida "verdadero" cuando las tres entradas son verdaderas. En este caso, las puertas "no" resultan innecesarias; el primer y segundo sumandos se aplican a una puerta de tipo "y", cuya salida se combina con el bit transportador en una segunda puerta tipo "y". Estos cuatro sistemas de puertas darán el bit suma correcto para todas las combinaciones de bits de entrada. Queda únicamente ligar las salidas de los cuatro grupos con puertas de tipo "o", de suerte que el sistema completo de puertas tenga una línea de sa-

lida única, la cual dé siempre el bit suma correcto.

La creación del bit transportador para la columna inmediatamente superior se obtiene de una forma más sencilla que el cálculo del bit suma. La tabla de verdad del bit transportador indica que el bit será 1 cuando dos, por lo menos, de los bits de entrada sean 1. Basta, pues, con combinar en pares las distintas entradas. Se destinan puertas de tipo "y" para el primer y segundo sumandos, para el primer sumando y el bit transportador (del orden inmediatamente inferior) y para el segundo sumando y el bit transportador. Las salidas de estas puertas de tipo "y" se combinan luego a través de una puerta tipo "o". Si dos, al menos, de las entradas son verdaderas, síguese que una al menos de las puertas de tipo "y" tendrá la salida verdad y se generará el bit transportador de valor 1.

Cuando se construye el sumador, hay tres terminales de entrada (para el primer sumando, para el segundo sumando y para el bit transportador de entrada),

A	B	C	BIT SUMA	BIT TRANS-PORTA-DOR
0	0	0	0	0
0	0	1	1	0
0	1	0	1	0
0	1	1	0	1
1	0	0	1	0
1	0	1	0	1
1	1	0	0	1
1	1	1	1	1

LAS TABLAS de verdad para la suma binaria dan los valores calculados del bit suma y el bit transportador para todos los posibles valores de la entrada. Las entradas son los dos bits que han de ser sumados *(A* y *B)* y el bit transportador *(C)* de la potencia de 2 inmediata inferior. Las reglas para la adición especifican que el bit suma es un 1, si exactamente una de las entradas es un 1 o si las tres son unos; de lo contrario el bit suma será 0. El bit transportador para la potencia de 2 inmediata superior es un 1 si por lo menos dos de las entradas son unos.

que se aplican a todas las puertas convenientes a la vez. Las dos salidas corresponden a la suma y al bit transportador para el orden inmediatamente superior. En este diseño el sumador completo posee tres puertas del tipo "no", 11 puertas del tipo "y" y cinco del tipo "o". Resulta muy fácil conectar circuitos de este tipo uno a continuación del otro para poder sumar números binarios de más de un bit. Basta con conectar la línea de salida del bit transportador de una etapa dada al terminal de entrada transportador de la etapa siguiente.

En un ordenador, los símbolos del álgebra de Boole encuentran su expresión en el soporte físico de transistores y otros dispositivos electrónicos. En un microordenador, esta transformación resulta particularmente elegante, ya que la lógica simbólica se plasma de una manera casi directa sobre la superficie de una pastilla de silicio.

Vamos a considerar las realizaciones electrónicas de las puertas lógicas en una única tecnología de semiconductores: la del transistor, de efecto de campo de óxido metálico semiconductor de canal *n* incrementado (o abreviadamente MOS *n*) [véase "Elementos del circuito microelectrónico, de James D. Meindl, en este mismo número]. Estos transistores utilizan, para su alimentación, una tensión positiva y conducen cuando se aplica una señal positiva a la puerta. Aquí sólo se estudiará el circuito de configuración más sencilla para cada puerta. Estas no son, necesariamente, las configuraciones mejores, pero los refinamientos añadidos a los circuitos reales del ordenador no cambian ninguno de sus principios básicos de funcionamiento.

Los tres elementos de un transistor del tipo MOS *n* son el surtidor, la puerta y el drenador. La puerta (que no hay que confundir con la puerta lógica) controla la corriente que fluye entre el surtidor y el drenador. Cuando se aplica a la puerta un potencial positivo de dos voltios aproximadamente, el transistor conduce; si la tensión se reduce a cero, cesa la conducción. De nuevo, dichas tensiones pueden calificarse como "altas" y "bajas"; aquí convendremos en identificar la tensión alta con un 1 en sistema binario o un "verdadero" en lógica.

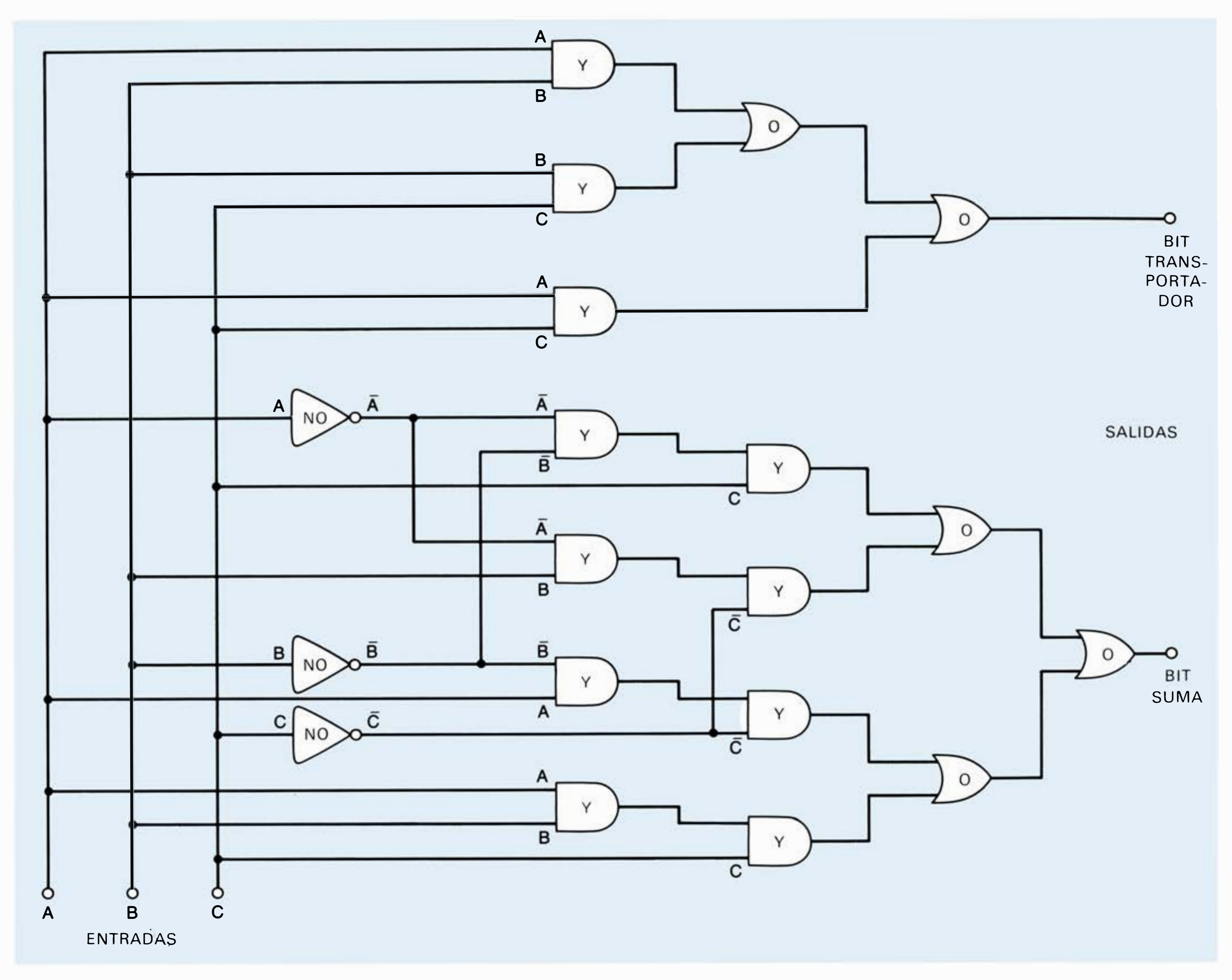
La más sencilla de las puertas lógicas es un inversor; éste desempeña la función lógica "no". Consta de un solo transistor que está conectado en serie con una resistencia. La tensión de alimentación (potencial positivo de unos cinco voltios) se aplica a la resistencia. La señal de entrada se aplica directamente a la puerta del transistor, y la salida se percibe en el punto de unión entre el transistor y la resistencia.

Cuando la entrada al inversor está en estado bajo, el transistor no conduce y, por consiguiente, no fluye corriente desde la alimentación hasta el punto de tierra, o retorno común del circuito, a través del resistor y el transistor. En consecuencia, no hay caída de tensión a través de la resistencia, y es como si la salida estuviese conectada a la tensión de alimentación. Cuando la entrada es baja la salida es alta.

Pero si ahora aplicamos a la puerta una señal alta, el transistor será conductor. Debido a la corriente que fluye por el circuito, la caída de tensión a través de la resistencia será notable. En este caso, la salida de la puerta está conectada de forma efectiva al punto de tierra, gracias a la baja resistencia que ofrece el transistor; se halla por tanto en estado de tensión baja. Tal es justamente el comportamiento requerido de una puerta del tipo "no".

Las otras puertas lógicas se construyen de acuerdo con principios similares. Para la puerta del tipo "y", que precisa dos entradas, se substituye el transistor único del inversor por una pareja de transistores conectados en serie. Las entradas son las puertas de los transistores en cuestión; por hallarse los transistores conectados en serie, síguese que sólo cuando ambas entradas sean altas quedará completado un camino conductor. De esta forma, el circuito actúa simultáneamente como una puerta del tipo "y" y como un inversor; a esta combinación de dispositivos lógicos se la denomina puerta del tipo "no-y". Pueden construirse dispositivos lógicos a partir de puertas de tipo "no-y", y, en ciertas tecnologías, ésta es la práctica usual; aunque el circuito puede convertirse fácilmente en la puerta normal de tipo "y". Todo lo que se requiere es poner, a continuación de la salida de la puerta tipo "no-y", un segundo inversor, el cual restablezca su propia polaridad a la señal. Cuando las dos entradas de la puerta sean altas, la salida de la parte del circuito constituido por la puerta "no-y" será baja; esto sitúa la salida del inversor –y del circuito en conjunto– en el estado de tensión alta.

Una puerta del tipo "o" se construye de la misma forma, con la salvedad de que los dos transistores de entrada, en vez de estar conectados en serie, lo están en paralelo. En virtud de este cambio, existe un camino conductor a tierra siempre que una de las dos entradas sea alta. Como en la puerta de tipo "y",



EL SUMADOR BINARIO ofrece una configuración de puertas lógicas para cada combinación de bits de entrada (*A, B, C*) que requiere una salida. Las cuatro filas de puertas de la parte inferior calculan el bit suma; las tres filas de la parte superior calculan el bit transportador para la potencia de 2 inmediata superior. A modo de ejemplo, supóngase que los dos bits que han de ser sumados (*A* y *B*) son ceros pero el bit transportador *(C)* de la columna precedente es un 1; la tabla de verdad *(figura anterior)* indica que esta combinación de entradas ha de producir un bit suma igual a 1 y un bit transportador igual a 0. El bit suma se genera en la cuarta fila, desde abajo, donde *A* y *B* se aplican a una puerta "y" en forma invertida, y así ambas aparecen como unos. La salida de esta puerta, que es un 1, se combina con el bit transportador en una segunda puerta "y", y la salida 1 pasa a través de una serie de puertas "o" hasta la salida del sumador. Ninguna puerta para el cálculo del bit transportador responde a esta combinación de entradas; por ello, el bit transportador es un 0; si dos o más entradas fuesen unos, se enviaría un bit transportador de valor 1. Un sumador binario requiere una configuración lógica como ésta para cada bit de los números que han de ser sumados, con la salida transportadora de cada etapa conectada con la entrada transportadora inmediata.

debe añadirse al circuito un inversor si a una entrada alta ha de seguir una salida alta. Sin el inversor, el circuito es una puerta "no-o".

Mediante la simple unión de puertas lógicas como las descritas, se puede construir un sumador binario de un bit a partir de 50 transistores y otras tantas resistencias. No es absolutamente necesario que un sumador sea tan complejo. Los refinamientos incorporados en el modelo para que un transistor desempeñe más de una función reducen el número de elementos a 17 transistores y cuatro resistencias, o simplemente a 21 transistores.

De todas formas, incluso en su expresión más sencilla, el circuito sumador es un aparato lógico de considerable complejidad. Por lo demás, sólo puede sumar números de un bit. Normalmente, los microprocesadores trabajan con números binarios de ocho a 16 bits; para un sumador completo se necesita un número correspondiente de réplicas.

Las manipulaciones lógicas y aritméticas de números binarios no son las únicas funciones de los circuitos de puerta fundamentales. El almacenaje de información constituye otra función básica del ordenador. Dicho almacenaje se lleva a cabo mediante la combinación de puertas para formar circuitos que reciben distintos nombres: basculadores, sujetadores o cerrojos, registros y células de memoria; la designación particular depende, sobre todo, de la misión a la que se destina el circuito.

Un pulso de tensión aplicado a la entrada de una puerta lógica tarda en pasar por ella unas cuantas decenas de nanosegundos (un nanosegundo es la mil millonésima parte de un segundo), sin dejar ninguna traza de su movimiento. Para un funcionamiento seguro, un complejo dispositivo de puertas ha de tener un sentido histórico. Tal es la función de un sujetador, a saber: se trata de un circuito que mantiene el estado de su salida indefinidamente como respuesta a una señal de entrada fugaz.

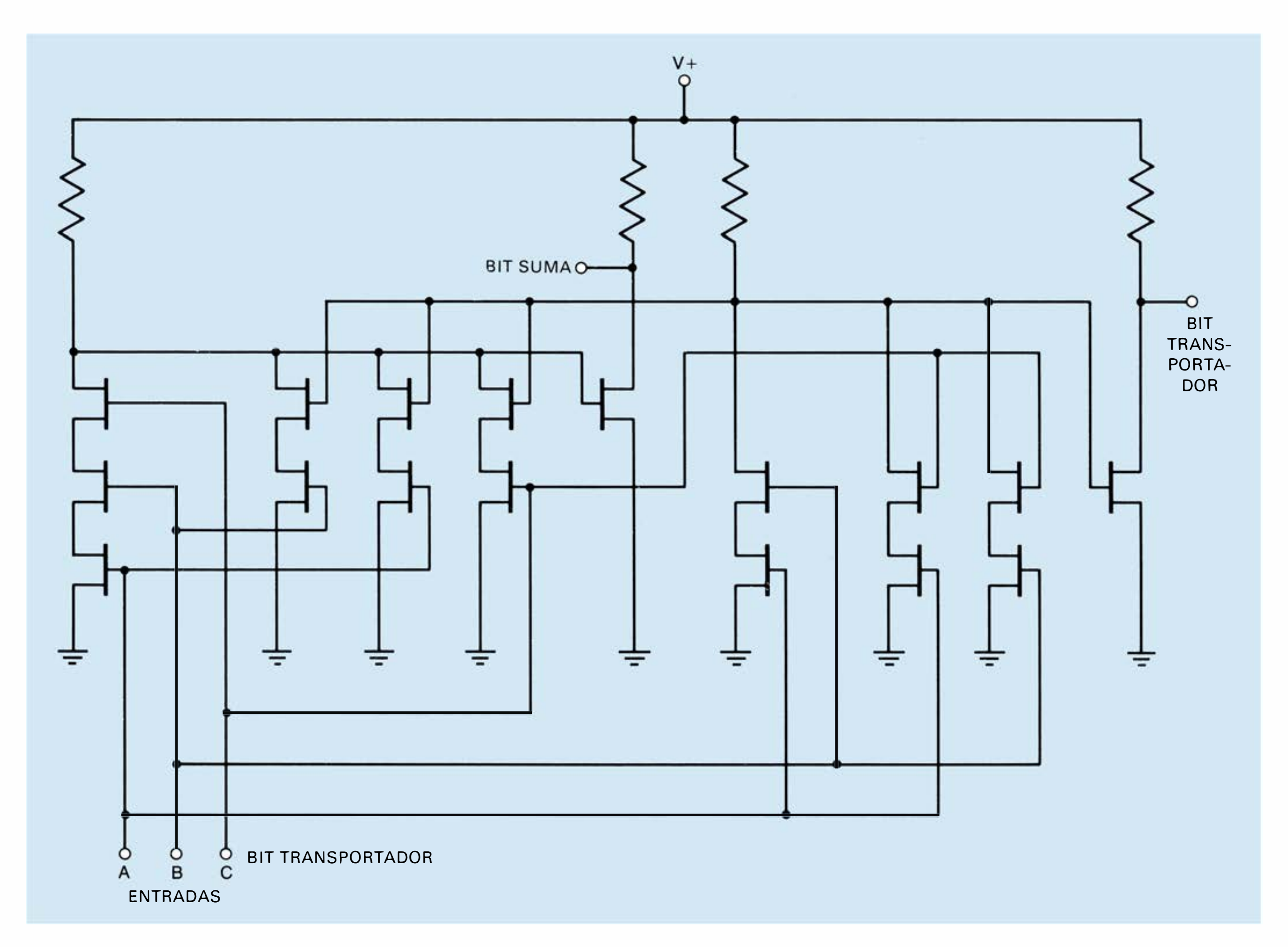
Con dos puertas de tipo "o" y dos inversores puede construirse un sujetador

simple. Una de las entradas de cada puerta del tipo "o" se reserva como entrada del circuito; las entradas restantes reciben la salida invertida de la puerta opuesta. Devolviendo las señales de salida invertidas a los terminales de entrada, se crea un proceso de realimentación positiva, el cual estabiliza el circuito en uno de los dos estados. Se denomina "posición" y "reposición" a las entradas al circuito. La salida será un cero si el último pulso recibido se aplicó a la línea de reposición; y será un 1, si el pulso recibido en último lugar vino por la línea de posición. Este circuito, conocido también como basculador o bioestable (se le designa con los términos ingleses "flip-flop", en razón de su forma de cambiar de estado), puede fabricarse con sólo dos transistores.

Un dispositivo algo más versátil es el llamado sujetador de registro. Puede construirse precediendo las entradas al basculador con un circuito simple, realizado con dos puertas de tipo "y" y un inversor. De las dos entradas a este circuito, una queda reservada para la señal de control y la otra recibe datos. Mientras la señal de control sea baja, los pulsos de datos se mantendrán bloqueados y no tendrán efecto alguno en el estado de la salida. Cuando la línea de control sea alta, la salida cambiará para reflejar el estado instantáneo de los datos, manteniendo dicho estado incluso después de que se haya eliminado el pulso de control. Un sujetador sólo puede almacenar un bit de información; de donde se sigue que un registro completo exige un sujetador por cada bit que tenga que almacenarse.

En los microprocesadores se emplean registros para un almacenamiento temporal de los datos, de los resultados parciales, de las instrucciones y de las direcciones de donde hay que buscar nuevos datos e instrucciones. Se ha diseñado un registro especial para calcular sin solución de continuidad, desde cero hasta su capacidad máxima; y se emplea universalmente para proceder a través de una secuencia de instrucciones.

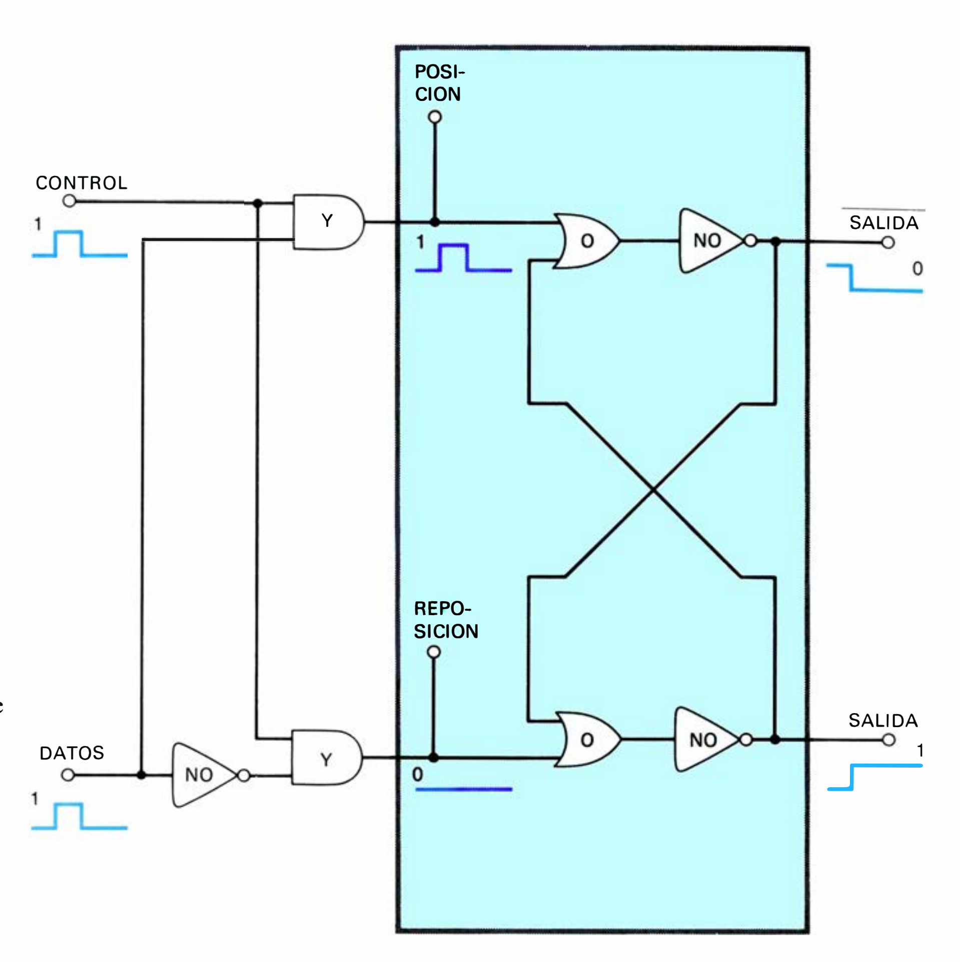
Para que un ordenador pueda funcionar se precisan bastantes más dispositivos lógicos. Así, un comparador determina si dos números binarios son o no iguales. Los bits particulares de los números sometidos a comparación se aplican a las entradas de puertas de tipo "y"; y las salidas de tales puertas se combinan de suerte que sólo se produzca un bit alto si todos los bits son idénticos. Un decodificador acepta un núme-



CIRCUITO SUMADOR, equivalente funcional de las formaciones lógicas. Su estructura está simplificada, porque algunos elementos desempeñan más de una función. El bit transportador se genera fácilmente mediante los siete transistores de la derecha. Si dos cualesquiera de las entradas se hallan en la fase de tensión alta, entonces uno de los tres pares de transistores puestos en serie conducirá; si uno al menos de estos pares conduce, entonces la salida común a los tres pares será forzada al estado de tensión baja. El transistor del extremo derecho invierte esta señal para producir un bit transportador alto, o un 1 binario. El resultado intermedio del cálculo del bit transportador se emplea también en el cálculo del bit suma. En la estructura de los siete transistores sitos a la izquierda, cada entrada se combina con la forma invertida del bit transportador, y la salida común de estas puertas es seguida por un inversor. De esta operación resulta que el bit suma es un 1 si uno de los bits de entrada es un 1 y el bit transportador es un 0; examinando la tabla de verdad para la suma resulta que este procedimiento siempre da la suma correcta. Los tres transistores en serie del extremo izquierdo generan un 1 de salida (a través del inversor), si las tres entradas son unos. El circuito emplea 17 transistores y 4 resistencias; en un circuito integrado éstas pueden ser substituidas por transistores.

ro binario como entrada y activa una salida, o en algunos casos más de una, en razón del valor del número. El decodificador que acepte un número de tres bits podría seleccionar una de entre ocho líneas, puesto que hay ocho números binarios de tres bits cada uno (que van desde 000 hasta 111).

Cualquier función expresable en términos de variables binarias puede quedar representada por un ordenamiento de puertas lógicas. Hemos visto cómo se construye un circuito para la suma echando mano de una docena de puertas, más o menos, por cada bit de los números sumandos. Aplicando los mismos principios pueden llevarse a cabo operaciones mucho más elaboradas. Podríamos idear, por ejemplo, una configuración lógica capaz de elevar un número al cuadrado, multiplicar luego el cuadrado por un coeficiente, sumar más tarde a este resultado el producto del número original con otro coeficiente y sumar, por último, una constante a ese gran total. Un circuito de ese género podría calcular expresiones cuadráticas. Incorporando, como mínimo, tres multiplicadores (cada uno de los cuales es a su vez un circuito complejo), amén de sumadores, registros y otros componentes, se convertiría en un dispositivo caro y complejo. Y lo que es más, sólo podría manipular expresiones cuadráticas; para afrontar problemas de ecuaciones cúbicas habría que idear otro circuito mucho más elaborado. Si un ordenador tuviese que exigir distintas configuraciones de puertas lógicas para cada función que hubiera de desarrollar, se convertiría en una máquina carente de versatilidad, desmesuradamente grande y prohibitivamente costosa.



LOS CIRCUITOS CERROJOS O BIESTABLES mantienen un registro de señales momentáneas. En su forma más sencilla *(recuadro coloreado)* el circuito tiene entradas designadas por "posición" y "reposición". Con la entrada de reposición en estado bajo, un breve pulso alto aplicado a la entrada "de posición" fuerza a la salida a ponerse en un estado de tensión alta; estado que se mantendrá a partir de entonces indefinidamente gracias a una realimentación proveniente de la salida. Con la línea de posición baja, un pulso alto aplicado al terminal de reposición coloca de nuevo la salida en cero. Un segundo terminal de salida suministra una señal invertida. Este circuito puede ir precedido por un sistema de puertas para convertirlo en un registro sujetador controlado. Una de las entradas recibe ahora el nombre "control"; ninguna señal puede alcanzar el basculador o biestable a menos que la señal de control sea alta. En presencia de una señal de control alta, los datos pueden aplicarse a la otra entrada. Todos los pulsos altos se dirigen a la entrada de posición y los bajos a la entrada de reposición; así pues, la salida siempre refleja el estado de la señal dato más reciente.

Si los ordenadores reales se las arreglan con pocos elementos lógicos es porque cambian eficazmente las interconexiones entre los dispositivos, creándose así cualquier configuración lógica pretendida. Un cálculo puede necesitar de varias sumas, pero el ordenador sólo exige un circuito sumador, que se va alimentando con cada nuevo conjunto de operandos. Para utilizar los elementos lógicos de esta manera flexible se necesita un sistema de distribución del tiempo y circuitos de control mucho más elaborados. Y se requiere algo más, algo que hasta el presente no hemos considerado en los circuitos lógicos vistos, a saber: un programa, esto es, un conjunto de órdenes para cambiar la configuración del entramado lógico.

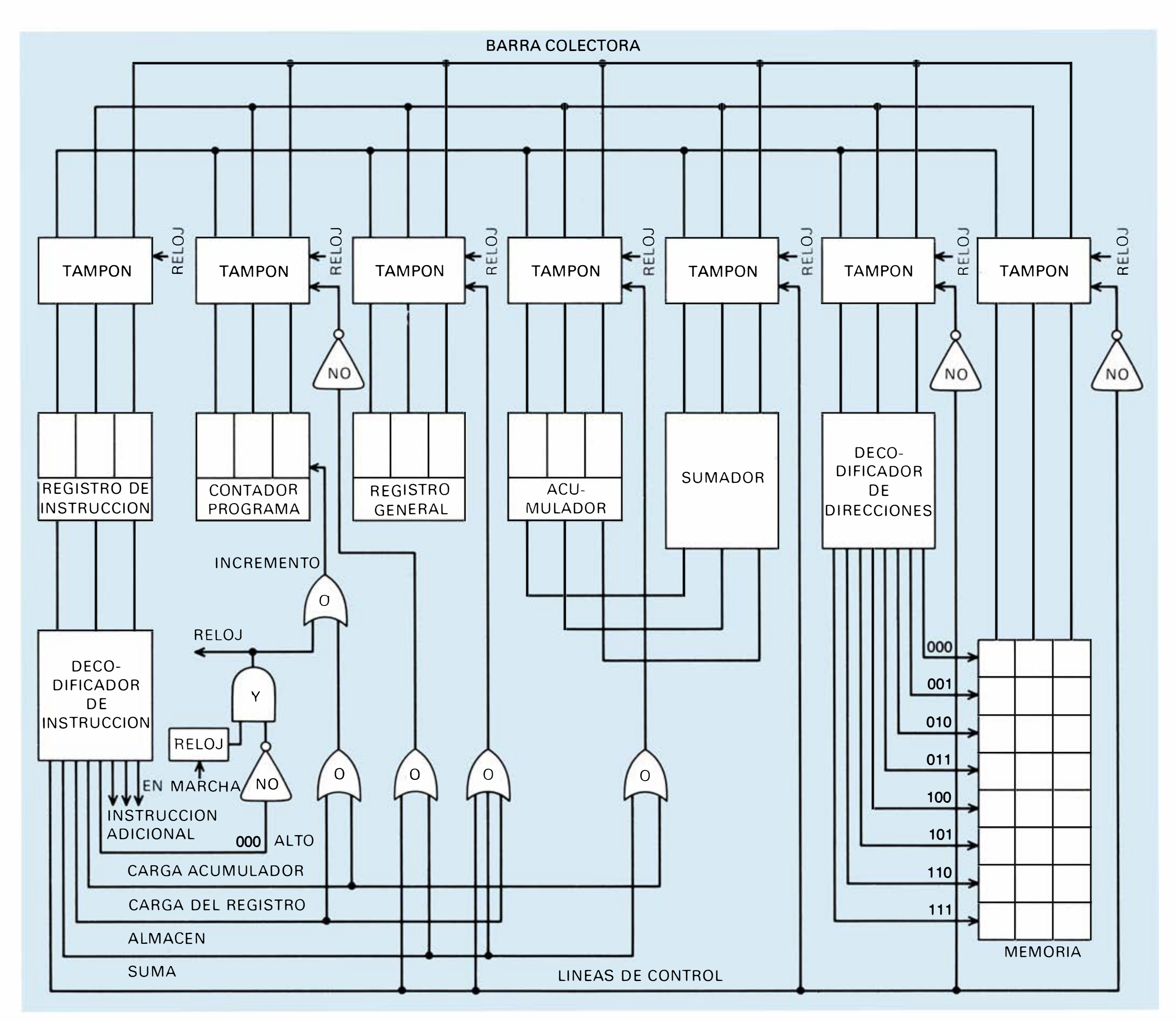
La idea de programa almacenado fue introducida por Charles Babbage en 1833, como peculiaridad de una "máquina calculadora" que propuso construir. Su ingenio, que hubiera empleado un engranaje mecánico para operar con aritmética decimal, habría incorporado bastantes principios que son fundamentales en el diseño de los ordenadores electrónicos. Aunque la idea era buena, la calculadora en cuestión no llegó nunca a construirse. La versión moderna del programa almacenado se debe a John von Neumann, el cual añadió una innovación importante. Determinó que tanto los datos sujetos a procesamiento como las instrucciones para procesar a éstos tenían que escribirse con la misma notación. Mediante esta estratagema, la máquina podía manejar las instrucciones como si fueran datos. Un programa puede, pues, modificar otro programa e incluso a sí mismo.

Vale la pena detenerse un momento para entender cómo los elementos lógicos de un ordenador se controlan mediante el programa. Para lo cual, inventaremos un hipotético microprocesador y un programa de suma de dos números binarios.

Un principio operacional de casi todos los microprocesadores es que las distintas partes de la máquina se comunican entre sí a través de un sistema de conductores denominados barra colectora. Esta consta sencillamente de un conjunto de vías conductoras que conectan entre sí distintas zonas de la pastilla del microprocesador, vías que se extienden a través de clavijas conectoras hasta un conjunto de conductores paralelos fuera de la pastilla.

Todos los dispositivos lógicos que configuran el ordenador se hallan conectados a la barra colectora, aunque no operen simultáneamente todos. La memo-

ria, en donde se almacenan datos y programa, comunica esa información a la máquina a través de la barra colectora. La unidad lógica y aritmética (ALU, del inglés "arithmetic and logic unit"), como reza su nombre, contiene dispositivos tales como los sumadores para operaciones aritméticas y lógicas sobre los datos recibidos por la barra colectora. A ésta se conectan varios registros para un almacenamiento temporal de los operandos y del resultado; de tales registros, el llamado acumulador se halla estrechamente asociado con la ALU. Un segundo registro suele tener por misión almacenar las direcciones de memoria en donde deben buscarse los datos y las instrucciones necesarias. Hay un tercer registro, denominado contador del programa, que encierra siempre la dirección de la instrucción inmediata que ha de realizarse. Por último, se encuentran unidos a la barra colectora los circuitos de decodificación y de control que, no obstante, pueden comunicar con los elementos de la máquina a través de líneas de control aisladas. A todas partes de la máquina llegan señales de reloj por vías conductoras que son independientes de la barra colectora.



UN HIPOTETICO ORDENADOR es una configuración de elementos lógicos electrónicos cuyas interconexiones pueden alterarse por instrucciones almacenadas en el mismo ordenador. Instrucciones, operandos y las "direcciones" que designan las localizaciones en la memoria del computador se expresan todas ellas por números binarios. En esta máquina, los números binarios son de tres bits; como sólo hay ocho combinaciones posibles con tres bits, los operandos sólo pueden tener ocho valores, la máquina puede tener tan sólo ocho localizaciones de memoria y está limitada a un repertorio de ocho instrucciones. La vía principal de comunicación en la máquina es la barra colectora, conjunto de conductores paralelos con conexiones con la mayoría de los dispositivos. Estos dispositivos comprenden cuatro registros: el contador del programa, que normalmente contiene la dirección de la instrucción siguiente; el registro de instrucciones, donde se retiene el código binario para cada operación; un registro general para el almacenamiento temporal de operandos o direcciones, y el acumulador, un registro con conexiones con la barra colectora y con la unidad aritmética y lógica (ALU) del ordenador. En este caso la ALU es un puro sumador, pero en las máquinas reales está capacitada para realizar varias operaciones. Los datos son escritos en la memoria o leídos a partir de ella en una localización especificada por el decodificador de direcciones. El decodificador consiste en una serie de puertas lógicas que activa una de las ocho líneas de direcciones en respuesta a un número binario de tres bits. Las instrucciones quedan decodificadas de idéntica manera. Cada uno de estos dispositivos está aislado de la barra colectora mediante un "tampón" que conecta el dispositivo con la barra colectora solamente cuando ha recibido una señal de control del decodificador de instrucciones, junto con una señal simultánea procedente del reloj del sistema; el reloj emite un tren continuo de pulsos a todas las partes de la máquina. La información que aparece en la barra colectora representa un operando, una instrucción o una dirección, según el estado de las señales de ritmo temporal y control. Las posibilidades de un ordenador tan pequeño son triviales; refleja principios de organización muy comunes a los de las máquinas reales, microprocesadores incluidos.

La barra colectora simplifica la comunicación en el interior del ordenador. En vez de instalar una línea particular desde cada parte de la máquina a cualquier otra, todos los dispositivos comparten una barra colectora, o varias. El tributo que hay que pagar por esta simplificación estriba en que los distintos dispositivos han de usar la barra colectora de uno en uno. Una vez más, se requieren circuitos de control y de distribución de tiempo más perfeccionados. Para ejercer el control, hay que desconectar eficazmente de la barra colectora todos los dispositivos, a excepción de los que en un determinado momento estén transmitiendo o recibiendo datos.

Cuando un microprocesador está bajo el control de un programa, oscila entre dos ciclos: el ciclo de aportación de las instrucciones y el ciclo de ejecución de las mismas. Aunque la distribución de los tiempos de los acontecimientos varía de un microprocesador a otro, asumiremos aquí que, en nuestra máquina, los ciclos se dividen a su vez, de suerte que la máquina tenga un total de cuatro etapas. En la primera, se envía una dirección a la memoria y, en la segunda, se remite, desde ella, una instrucción. En la tercera etapa, se envía otra dirección a la memoria; en la cuarta, de la memoria se remite un operando, el cual se procesa de acuerdo con la instrucción recibida.

Partiremos de suponer que el programa está convenientemente almacenado en la memoria, empezando en la posición cero y que el contador de programa se ha repuesto a cero. Cuando la máquina se pone en marcha, entra inmediatamente en la primera etapa: el contenido del contador de programa se sitúa en la barra colectora y se interpreta como una dirección de la memoria. Se puede seguir este procedimiento de partida, gracias a que el reloj, que define las cuatro etapas, controla un sistema de puertas que permite que sólo el contador de programa transmita señales a la barra colectora durante la primera etapa. Mientras dura este período, sólo el sistema de memoria está capacitado para recibir señales procedentes de la barra colectora.

La dirección emitida por el contador del programa es cero, ya que hemos dado por supuesto que dicho contador se había repuesto a este valor. Este número binario ha de ser decodificado por puertas que forman parte del sistema de memoria, de suerte que se aplique una señal a las células de memoria con la dirección cero. El número de bits de la dirección determina el número de localizaciones de memoria máximo. Una dirección de ocho bits puede seleccionar solamente entre 256 lugares, puesto que sólo hay 256 números binarios de ocho bits cada uno. Una dirección de 16 bits da acceso a 65.536 posiciones de memoria. (Pueden especificarse posiciones adicionales de memoria si se aportan señales de control suplementarias.)

La señal aportada por un circuito de decodificación de direcciones a la situación o localización 0 es una señal de lectura; en respuesta a ella, el contenido de esa situación de memoria aparece en las líneas de salida de la memoria. Cuando la máquina abandona la primera etapa y entra en la segunda, el contador del programa y el decodificador de dirección se desconectan de la barra colectora. Al mismo tiempo, la salida de la memoria se conecta a ella, una vez más a través de las puertas controladas por el reloj del sistema. Los contenidos de la localización cero de la memoria aparecen, pues, en la barra colectora durante el segundo período marcado por el reloj. Se recordará que las señales presentes en la barra colectora durante este período, han de interpretarse como una instrucción; de esta forma, el reloj tendrá que activar también los circuitos de decodificación y control de la instrucción. En la unidad de control, la instrucción queda momentáneamente almacenada en un registro.

Esta secuencia de sucesos ha servido para traer una instrucción de la memoria y cargarla en el registro de instrucciones. Los contenidos de ese registro determinarán luego lo que ocurra en el resto del ciclo de la máquina. La decodificación de la instrucción es un paso crucial en las operaciones de la máquina. Como en el caso de la decodificación de las direcciones de memoria, el número de bits determina el número máximo de instrucciones, pero rara vez esto constituye una limitación seria: son pocos los microprocesadores que poseen más de 100 instrucciones en su repertorio.

La decodificación de instrucciones puede llevarse a cabo por varios caminos. Hay un método que emplea un decodificador "rígido"; en él, una configuración de puertas selecciona, para cada combinación de bits del código en que se opera, una única combinación de líneas de salida activas. Estas salidas son las líneas de control que conectan y desconectan, de la barra colectora, varios dispositivos; realizan asimismo otras funciones, verbigracia, aumentan los contadores e inician los cálculos en la ALU. Las instrucciones pueden deco-

dificarse también aplicándolas a una memoria sólo de lectura, donde queda permanentemente almacenado un modelo fijo de bits, una vez construido el procesador. Cada código de las operaciones se interpreta entonces como una dirección en la memoria sólo de lectura; se registra la salida de la memoria y se aplica a las líneas de control.

Supondremos convencionalmente, que la instrucción obtenida de la localización o situación 0 de memoria es un número binario al que se le ha asignado el significado: "busca en la memoria la próxima palabra y cárgala en el acumulador". Esta instrucción se ejecutará de la siguiente manera. El decodificador de instrucciones (sea cual fuere su mecanismo) puede construirse de manera que el conjunto de bits que represente tal instrucción active cinco líneas de control. Una de estas líneas de control incrementa el contador del programa, de suerte que éste contendrá ahora la dirección de la próxima localización: 1. Las otras cuatro hacen posible que el contador del programa, el acumulador, el decodificador de direcciones y las líneas de salida de la memoria se conecten a la barra colectora. Pero todos estos dispositivos no se conectan a la vez, aun cuando todos ellos reciban las señales de control simultáneamente. Cuando el reloj marca el tercer período, indicando el comienzo del ciclo constituido por la ejecución de la instrucción, la máquina entra en una fase reservada para enviar una dirección a la memoria. Así pues, sólo la fuente de la dirección (en este caso, el contador del programa) y su destino (el decodificador de dirección) son activados durante este período por una señal del reloj y por una señal de control.

En la cuarta fase del reloj, el contador del programa y el decodificador de dirección quedan aislados de la barra colectora, permaneciendo conectadas a la misma las salidas de la memoria y el acumulador. De lo cual resulta que el número almacenado en la localización o situación de memoria 1 se carga en el acumulador.

En este momento, el ordenador ha completado un ciclo de máquina entero. El contador de programa se ha incrementado de nuevo automáticamente, de forma que ahora contiene el número 2, que en notación binaria se escribe 010; aquí es donde debe buscarse la instrucción inmediata. Cuando el reloj dé una nueva señal, el ordenador vuelve a su estado inicial y el ciclo se repite. En primer lugar, los contenidos del contador del programa se colocan en la barra colectora y se interpretan como una dirección; luego, en la segunda fase del reloj, los contenidos de la localización de memoria direccionada se interpretan como una instrucción. Esta segunda instrucción puede reclamar que los contenidos de uno de los registros se consideren como una dirección de memoria durante la tercera fase del reloj; los datos así obtenidos se cargarán en otro registro, a lo largo de la cuarta etapa. El segundo ciclo de la máquina quedaría así cerrado y el contador de programas aumentaría de nuevo (a 3, que, en números binarios, sería 011).

La instrucción operante durante la primera parte del siguiente ciclo de la máquina podría exigir que se sumasen los contenidos de un registro a los del acumulador. En cuyo caso, no hay necesidad de buscar en la memoria un operando durante la tercera fase del reloj; de esta suerte, tal etapa podría quedar ignorada,

CICLO DE LA MAQUINA	PERIODO RELOJ	CONTADOR DEL PROGRAMA	INFORMACION AL INTERCOMUNICADOR		SIGNIFICADO DE LAS INSTRUCCIONES
1 BUSQUEDA	1	000	CONTENIDOS DEL CONTADOR DEL PROGRAMA	000	CARGA LOS CONTENIDOS DE LA PROXIMA LOCALIZACION DE MEMORIA EN EL REGISTRO GENERAL
	2	000	INSTRUCCION "CARGA EL REGISTRO"	010	
EJECUCION	3	001	CONTENIDOS DEL CONTADOR DEL PROGRAMA	001	
	4	001	DATOS	111	
2 BUSQUEDA	1	010	CONTENIDOS DEL CONTADOR DEL PROGRAMA	010	CARGA LOS CONTENIDOS DE LA PROXIMA LOCALIZACION DE MEMORIA EN EL ACUMULADOR
	2	010	INSTRUCCION "CARGA EL ACUMULADOR"	001	
EJECUCION	3	011	CONTENIDOS DEL CONTADOR DEL PROGRAMA	011	
	4	011	DATOS	001	
3 BUSQUEDA	1	100	CONTENIDOS DEL CONTADOR DEL PROGRAMA	100	SUMA LOS CONTENIDOS DEL REGISTRO GENERAL A LOS CONTENIDOS DEL ACUMULADOR Y RETIENE EL RESULTADO EN EL ACUMULADOR
	2	100	INSTRUCCION "SUMA"	100	
EJECUCION	3	100	(DESOCUPADO)	—	
	4	100	CONTENIDOS DEL REGISTRO GENERAL	111	
4 BUSQUEDA	1	101	CONTENIDOS DEL CONTADOR DEL PROGRAMA	101	ALMACENA LOS CONTENIDOS DEL ACUMULADOR EN LA MEMORIA EN UNA DIRECCION ESPECIFICADA POR LOS CONTENIDOS DEL REGISTRO GENERAL
	2	101	INSTRUCCION "ALMACENA"	011	
EJECUCION	3	101	CONTENIDOS DEL REGISTRO GENERAL	111	
	4	101	CONTENIDOS DEL ACUMULADOR	000	
5 BUSQUEDA	1	110	CONTENIDOS DEL CONTADOR DEL PROGRAMA	110	PARA TODAS LAS OPERACIONES
	2	110	INSTRUCCION "ALTO"	000	
EJECUCION	3	110	—	—	
	4	110	—	—	

PROGRAMA de un ordenador de tres bits para la suma de dos números. El reloj del sistema divide a cada ciclo de la máquina en cuatro períodos; los dos primeros están dedicados a la búsqueda de una instrucción y los dos últimos a ejecutarla. En el primer período del reloj los contenidos del contador del programa (inicialmente 000) se colocan en la barra colectora y se interpretan como una dirección de la memoria; en el segundo período, los contenidos de la localización de memoria designada se almacenan en el registro de instrucciones. Los acontecimientos del tercer y cuarto períodos dependen de las operaciones especificadas por la instrucción. En el primer ciclo de la máquina, el código de la operación es 010, al que se le ha asignado el significado "busca los contenidos de la localización siguiente de la memoria y carga dicho número en el registro general". En respuesta a este modelo, el decodificador de instrucciones activa una línea de control que incrementa el contador del programa, de suerte que el contador señala al siguiente "item" y capacita al registro general para recibir al valor almacenado. En el segundo ciclo de la máquina, se carga otro operando en el acumulador; luego, en el tercer ciclo, se suma el número del registro al del acumulador. Los números sumados son el 111 y el 001, y como la máquina no tiene recursos para el desbordamiento de capacidad ("overflow") el resultado es 000. En el cuarto ciclo, el valor del registro general, que era inicialmente un operando, se trata como una dirección.

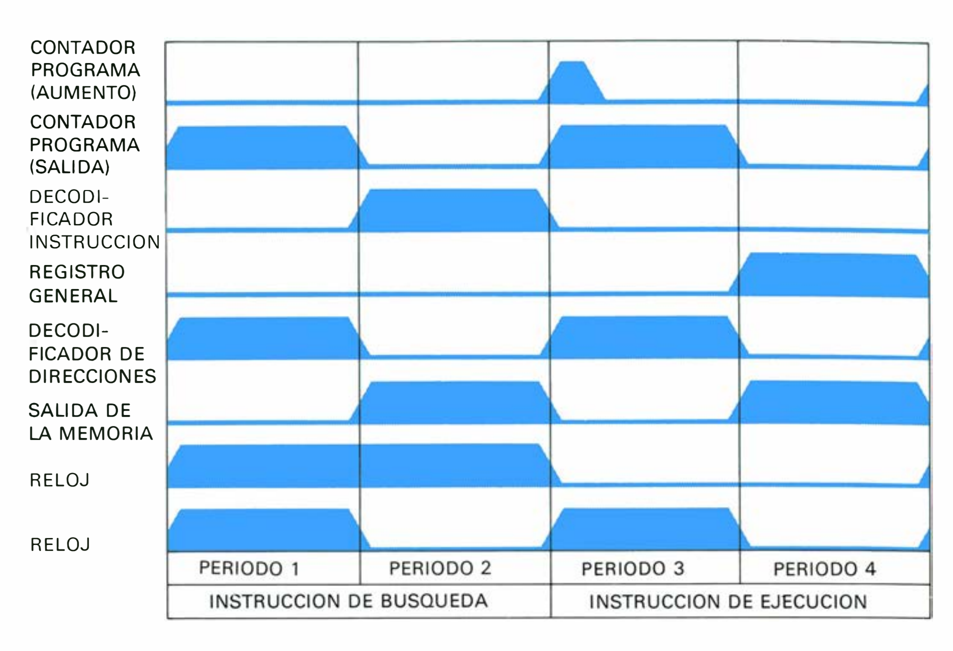
si no borrada, del proceso. Para ejecutar la instrucción, el circuito de control tiene que activar al sumador de la ALU y disponer que los operandos sean suministrados por los registros adecuados. Los resultados de las operaciones realizadas dentro de la ALU se retienen normalmente en el acumulador.

Instrucciones posteriores podrían determinar que se compare la suma obtenida por este método con un valor almacenado en la memoria. Si los dos números fueran diferentes, no se haría nada especial y la máquina pasaría a la siguiente instrucción de la secuencia. Si los números fuesen idénticos, se mandaría una señal que colocaría al contador del programa en cierto valor preseleccionado. La ejecución del programa se reanudaría en esta nueva localización de la memoria. Las operaciones de este estilo se llaman instrucciones de bifurcación condicional, dado que posibilitan que la máquina pueda elegir una de dos vías de un programa en función del resultado de un cálculo.

Otros códigos operativos pueden dar instrucciones de que la máquina lea datos procedentes de algunos dispositivos externos y periféricos, o bien de que la máquina suministre datos de salida a dichos dispositivos. Al final, la máquina deberá llegar a una instrucción de parada, que detenga todas las operaciones.

La máquina que acabamos de describir no corresponde a ningún microordenador en particular, pero incorpora características comunes a muchos de ellos (incluso comunes a muchos de los grandes ordenadores). La búsqueda y ejecución alternativa de instrucciones, por ejemplo, es un modo de actuar casi universal. Un rasgo más fundamental es el uso de un mismo lenguaje –los números binarios– para todas las señales del sistema. Un grupo de bits que aparezca en la barra colectora de nuestro microprocesador puede representar una dirección en la memoria, una instrucción o un operando. Su interpretación depende de la fase de la máquina y, por consiguiente, depende críticamente de las señales de temporización y de control que, a su vez, están determinadas por los códigos de instrucción.

La integración a gran escala ha transformado al ordenador; de ser una máquina que tenía que construirse a partir de muchos componentes ha pasado a ser un componente que puede ser incorporado en sistemas mayores. Gran número de los microprocesadores existentes en el mercado encierran en una sola pastilla todos los circuitos que hemos descrito, con la excepción de la memoria de acceso aleatorio y, en muchos casos, del reloj (véase "Microprocesadores", de Hoo-Min D. Toong, en este número). Ciertos dispositivos de pastilla única comprenden incluso una cantidad de memoria limitada.



EL DIAGRAMA DE TIEMPOS da la secuencia de los acontecimientos, durante el primer ciclo de la máquina, en la ejecución del programa de la ilustración anterior. Dos señales del reloj dividen el ciclo de la máquina en cuatro períodos del reloj. En el primer período sólo el contador del programa y el decodificador de direcciones se conectan a la barra colectora. En el segundo lo son la salida de memoria y el decodificador de instrucciones. En el tercero y cuarto períodos, la máquina está bajo el control del reloj y del decodificador de instrucciones. Ambos juntos incrementan el contador del programa, al que luego conectan con el decodificador de direcciones a la barra colectora. En el cuarto período, la salida de la memoria y el registro general están activos. Por último, al regresar el reloj a su primer estado, el contador del programa aumenta de nuevo.

La primera generación comercial de dispositivos microelectrónicos se incluye hoy entre los circuitos integrados a pequeña escala. Constaban de pocas puertas. Incluso en los casos en que había múltiples puertas en una sola cápsula, no siempre se hallaban internamente conectadas; el circuito que definía una configuración lógica tenía que establecerse mediante conductores externos.

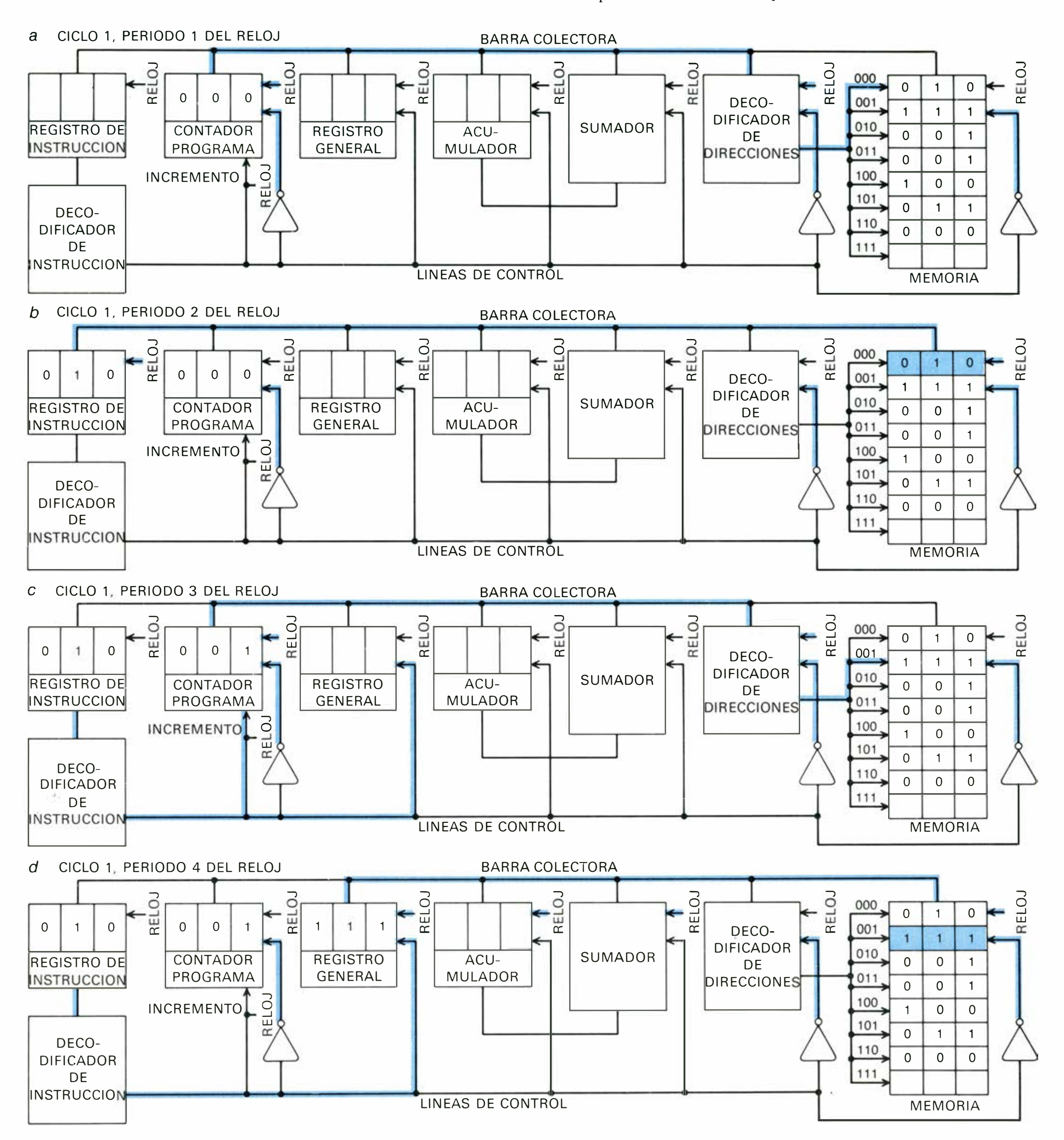
Los dispositivos que contienen más de 10 puertas en una pastilla, y menos de 200, constituyen los circuitos integrados a escala media. Muchas de estas pastillas contienen bloques funcionales completos de un ordenador. Por ejemplo, se pueden obtener pastillas que contengan registros, decodificadores, comparadores y contadores completos. Quizá la cota superior de la tecnología de los circuitos integrados a media escala esté marcada por las pastillas provistas de una unidad aritmética y lógica completa. Esta unidad acepta, por entrada, a dos operandos; además, puede realizar unas doce operaciones con ellos, de acuerdo con la fase en que se hallen las distintas líneas de control. Entre esas operaciones deben citarse la adición, la substracción, las operaciones lógicas "y" y "o" y la de desplazar un bit hacia la derecha o hacia la izquierda.

La mayor variedad de circuitos integrados a media escala se da en la tecnología del semiconductor bipolar llamada TTL ("transistor-transistor logic" = lógica transistor-transistor). Ni siquiera el desarrollo experimentado por los circuitos integrados a gran escala ha desplazado la utilidad de los dispositivos TTL. Muchos microprocesadores se han diseñado de tal forma que sus tensiones de señal y sus requerimientos de alimentación resulten compatibles con los circuitos TTL. Así pues, los dispositivos TTL pueden emplearse como componentes auxiliares de un sistema de microordenador para el control de la memoria o de los terminales de entrada y salida y para otras funciones, como, por ejemplo, generar señales de temporización.

La mayoría de los circuitos integrados a gran escala se han fabricado con transistores de óxido metálico semiconductor, y no con transistores bipolares de TTL; aquéllos pueden llegar a tener menores dimensiones y, además, empaquetarse de modo más compacto. El cambio de escala puede medirse a bulto contando el número de transistores que pueden incluirse en una pastilla. Los circuitos integrados a pequeña escala tienen del orden de 10 transistores en cada pastilla; los circuitos a media escala tienen del orden de 100. Los microprocesadores

fabricados recientemente a través de la integración a gran escala tienen entre 10.000 y 20.000 transistores. En la configuración geométrica regular de los elementos de las pastillas de memoria de semiconductores, el número de transistores se está acercando a los 100.000 (véase "Memorias microelectrónicas", de David A. Hodges, en este mismo número).

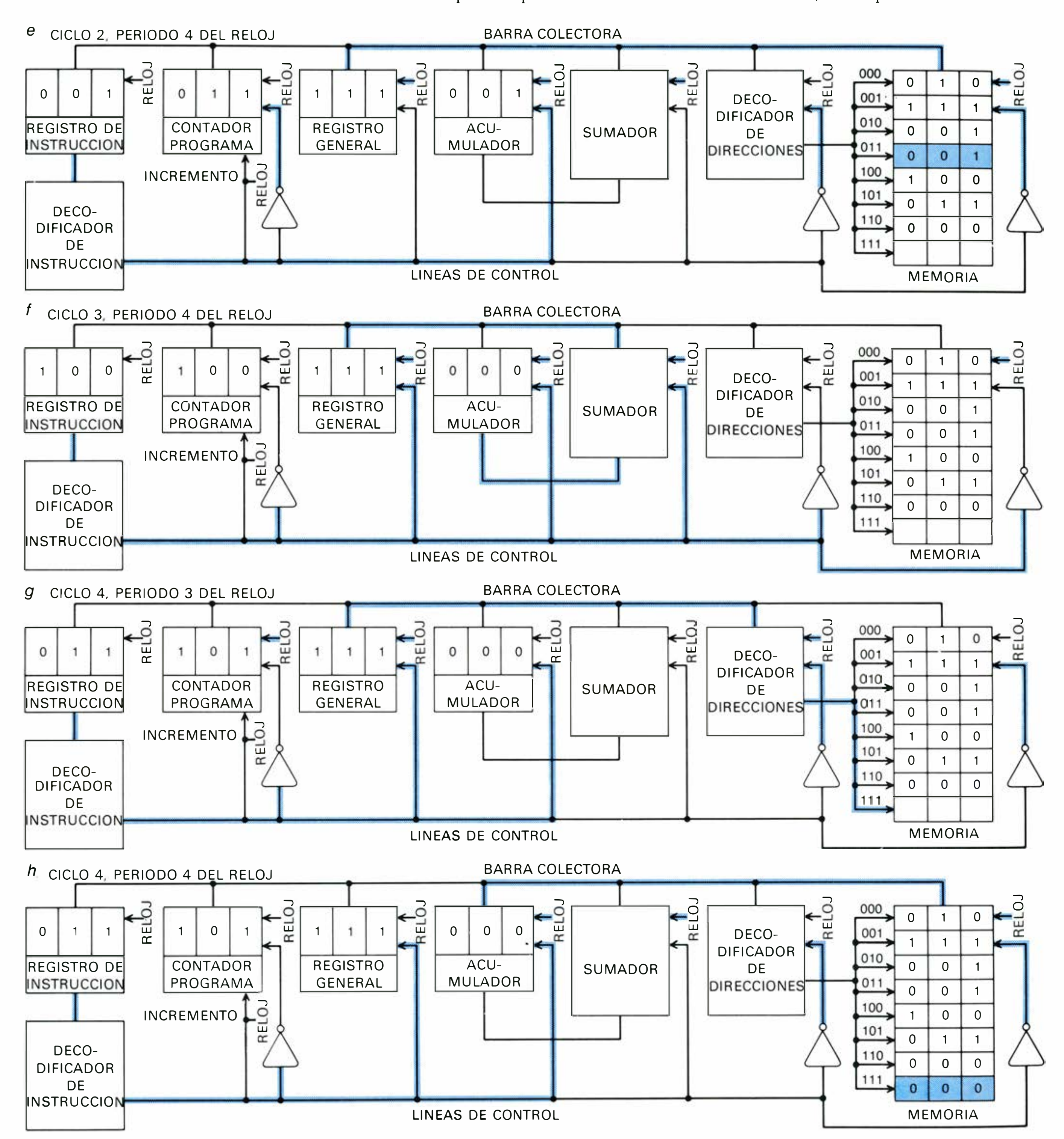
Tres son los factores que han intervenido en el rápido aumento del número de elementos del circuito por pastilla. Uno de los factores es el perfeccionamiento de las técnicas para desarrollar grandes cristales de silicio puro. Al aumentar el diámetro de las obleas –discos de silicio sobre los cuales se fabrican las pastillas–, pueden construirse más pastillas al mismo tiempo, y ello reduce el coste por unidad. Además, se ha me-



EJECUCION DE UN PROGRAMA por ordenador de tres bits. Se desarrolla en una secuencia de diagramas simplificados. Inicialmente, el programa es almacenado en las primeras siete localizaciones de la memoria y el contador del programa se coloca en 000. Durante los dos primeros períodos del reloj, no se ha decodificado ninguna instrucción; por consiguiente, la máquina ha de actuar absolutamente bajo control del reloj del sistema. Este modo de operar se ha hecho posible mediante inversores en las líneas de control del contador de programa, el decodificador de direcciones y la salida de la memoria; así queda asegurado el que estos dispositivos sean activados por el reloj en ausencia de la señal de control. En el primer período del reloj *(a)*, el contador del programa y el decodificador de direcciones están conectados a la barra colectora; la posición 000 de la memoria queda designada para ser leída. Los contenidos de esta localización se cargan en el registro de instrucción durante el segundo período *(b)*. Hasta el tercer período del reloj las señales de control del decodificador de instrucciones no están disponibles. La línea de control, activada inmediatamente, incrementa el contador del programa y suministra una señal que posibilita que el registro general entre en conexión con la barra colectora; además, esta línea de control hace que no se desconecten el contador del pro-

jorado también la calidad de los materiales empleados, menguando consiguientemente el número de defectos por oblea. Lo cual comporta el aumento del tamaño máximo práctico de la pastilla, pues reduce la probabilidad de hallar un defecto en una zona determinada. El tamaño de una pastilla de un circuito integrado a gran escala ha aumentado de menos de 0,1 a 0,5 cm².

Un segundo factor es el desarrollo de la litografía óptica, proceso que transfiere a una superficie de silicio a todos los modelos que componen un circuito. Desarrollando los sistemas ópticos capaces de resolver estructuras más finas, el tamaño de un transistor típico, medido por la longitud de una puerta, ha pasado de los pocos milímetros que medía en 1965 a las 5 décimas de milímetro de hoy día. Finalmente, los depurados avances en la



grama, el decodificador de direcciones ni la salida de la memoria. Sin embargo, no todos estos dispositivos se activan a la vez. Dado que el tercer período del reloj se reserva para emitir la dirección de un operando a la memoria, el reloj activa únicamente al contador del programa y al decodificador de direcciones, durante este período. En el cuarto período *(d)* el contador del programa y el decodificador de direcciones se paran por mediación del reloj, activándose la salida de la memoria y el registro general. La instrucción decodificada durante el segundo ciclo *(e)* reclama otro valor de la memoria para que sea cargado en el acumulador. La adición se realiza durante el tercer ciclo *(f)*. Comoquiera que esta operación no precisa acceso a la memoria, se emiten señales por el decodificador de instrucciones que bloquean al decodificador de direcciones y a la salida de la memoria; el contador del programa queda también inutilizado. Al mismo tiempo, se activan el registro general, el acumulador y el sumador. En el tercer período del reloj del cuarto ciclo *(g)* los contenidos del registro general se colocan en la barra colectora; puesto que el único dispositivo que resta activo es el decodificador de direcciones, este valor se interpreta como una dirección en la memoria. Los contenidos del acumulador, en donde estaba retenido el resultado de la adición, se almacenan en esta localización durante el cuarto período *(h)*.

estructura de los circuitos que permiten un uso más eficaz de la superficie del silicio han conducido a multiplicar por cien la densidad de los transistores que posee una pastilla.

La incesante evolución del microordenador irá exigiendo nuevos aumentos en la densidad de empaquetamiento. A medida que se va alcanzando el límite de la resolución óptica, se advierte la necesidad de nuevas técnicas de litografía y de fabricación. Los patrones de los circuitos tendrán que hacerse con radiaciones cuya longitud de onda sea menor que la de la luz; por consiguiente, se reclamarán nuevas técnicas de fabricación, capaces de una definición mayor.

La escala actual de las estructuras microelectrónicas requiere una acuidad de borde de 2500 angstroms, que es aproximadamente la longitud de onda de la radiación ultravioleta empleada en la fabricación de los circuitos. Para formar imágenes nítidas de estructuras aún menores, se puede emplear, en la confección del molde, un haz electrónico dirigido. La mayor resolución potencial viene ofrecida por los sistemas litográficos de rayos X, cuya longitud de onda eficaz es de unos 10 angstroms.

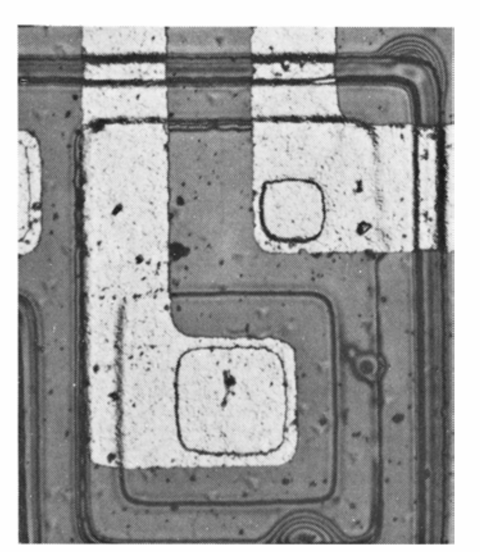
Moldeando con haces electrónicos y con rayos X, se pueden reducir las dimensiones lineales de un transistor en un factor de 50. Por tanto, la superficie ocupada por un transistor se reduciría más de 1000 veces. Si las virtualidades de la actual generación de microprocesadores resultan ya notables, las perspectivas para la próxima generación, las de los circuitos integrados a grandísima escala (VLSI, del inglés "very-large-scale integrated circuits") se sospechan extraordinarias. El procesador sólo ocupará una parte muy pequeña de tal pastilla, en la que podrían incluirse también millones de bits de memoria.

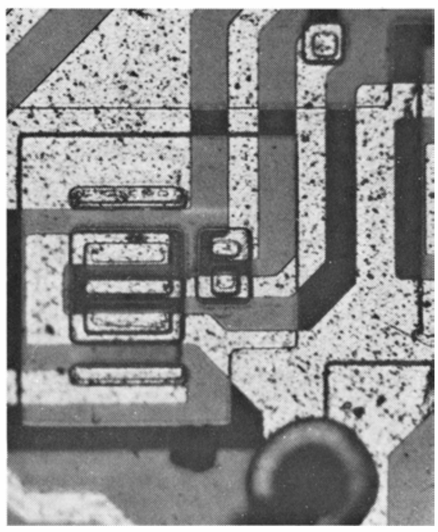
Aunque los obstáculos inmediatos con que nos topamos en la construcción de microprocesadores más complejos radican en los procesos de fabricación, su diseño tampoco es fácil. El proceso de diseño comienza con la determinación de la organización global del ordenador. La "anchura" de la máquina, o, en otras palabras, el número de bits de una palabra normalizada, es un aspecto fundamental al que hay que atender; y también, la cantidad de memoria sujeta a direccionamiento, que dictará el número de bits de dirección que han de suministrarse. Además de otras consideraciones múltiples, hay que prestar atención especial al número de conexiones a realizar con la pastilla; cada conexión requiere un soporte metálico situado en el borde de la pastilla, y solamente se dispone de un espacio reducido para dichos soportes.

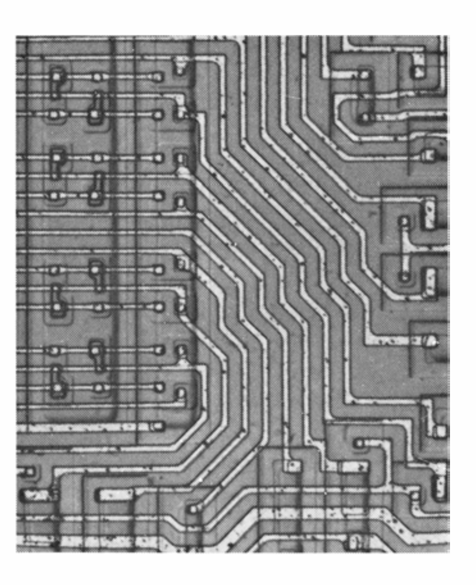
Decidida cuál será la arquitectura u organización de la máquina, se han de preparar diagramas lógicos para todas las funciones que tengan que desempeñarse. Los diagramas lógicos se reducirán luego a sus equivalentes electrónicos; por último, los diagramas de circuito resultantes han de expresarse en términos de dispositivos electrónicos y vías de intercomunicación entre ellos, que puedan fabricarse a un precio razonable dentro de la superficie de silicio asignada (véase "La fabricación de los circuitos microelectrónicos", de William G. Oldham, en este mismo número).

Las propiedades de los transistores individuales y de los otros componentes dependen substancialmente de su tamaño. Si el circuito ha de ser fiable, las dimensiones se mantendrán dentro de los límites prescritos, a pesar de las variaciones del proceso de fabricación. La temporización importa a su vez muchísimo: cuando se emite una señal por una parte del ordenador, otra parte ha de estar lista para recibirla. Una manera de atacar estos problemas consiste en la construcción de un prototipo del procesador, utilizando los componentes de que se dispone. En los últimos años se ha hecho mucho más común la construcción de prototipos con ayuda de un ordenador. Se construyen y combinan modelos matemáticos de distintos circuitos elementales para formar modelos de sistemas enteros.

Cuando está terminada la fase de diseño, el propio ordenador puede preparar los dibujos plantilla que se emplearán para definir la configuración del circuito sobre el silicio. Otros ordenadores se encargarán de supervisar el proceso de fabricación, y otros terceros comprobarán las pastillas terminadas, marcando automáticamente las defectuosas. No hay que gastar más palabras en demostrar la necesidad del desarrollo evolutivo en la creación de una tecnología compleja. Los circuitos integrados que se diseñan y construyen actualmente no hubieran sido posibles sin la asistencia de los realizados en los últimos años.







LA DISMINUCION DE ESCALA de los dispositivos electrónicos aparece en estas tres pastillas vistas con el mismo aumento: unas 300 veces. A la izquierda, una pastilla fabricada en 1961; un único transistor llena por completo todo el campo de visión. La pastilla del centro se fabricó en 1965; pueden distinguirse porciones de algunos elementos activos. El circuito integrado a gran escala de la derecha se fabricó en 1975; engloba más de doce transistores en la misma superficie de silicio. Todos estos dispositivos emplean la tecnología del semiconductor bipolar.

Ciencia y sociedad

Einstein contra Bohr

Del 24 al 29 de octubre de 1927, se reunía en Bruselas, bajo los auspicios del Instituto Solvay y en sus locales, el quinto Consejo Internacional de Física. La conferencia versó sobre el tema "Electrones y fotones". A la misma asistieron eminentes físicos de la época: H. A. Lorentz, que presidía las reuniones, N. Bohr, M. Born, P. Ehrenfest, A. Einstein, L. de Broglie, E. Schrödinger, W. Heisenberg y otros. Aunque establecidos ya claramente los postulados de la nueva mecánica cuántica, puede afirmarse que fue en esta reunión cuando la comunidad científica internacional le dio su voto de confianza. Quedaban muchos problemas por resolver. Nada tiene de extraño, pues, que entre los presentes, y muy particularmente entre Einstein y Bohr, se suscitaran vivas discusiones.

Una de las principales cuestiones a debatir era la renuncia a la descripción determinista que implicaban los nuevos métodos introducidos en la mecánica cuántica. En concreto, se estudió hasta qué punto la mecánica de ondas ofrecía posibilidades de un alejamiento menos radical de las ideas deterministas. El propio Lorentz, cuya labor moderadora fue muy destacada, inició los debates con las siguientes palabras: "¿No podríamos acaso conservar el determinismo convirtiéndolo en objeto de una creencia? ¿Es realmente necesario hacer del indeterminismo un principio?".

Uno de los puntos en el que se apoyaban los defensores de la teoría, como importante baza a su favor, era el carácter estadístico de la interpretación de la experiencia de la física mediante imágenes de ondas, que había tratado con éxito los problemas de colisiones.

Fue apareciendo también, cada vez con mayor intensidad, el carácter simbólico de la nueva concepción en la substitución del sistema ordinario de coordenadas espaciales de tres dimensiones por una representación del estado de un sistema con varias partículas mediante funciones de onda en un espacio de configuración; en este caso, las coordenadas eran tantas cuantos grados de libertad tuviera el sistema. En el curso de las discusiones, el último punto se relacionó con los avances obtenidos en el tratamiento de sistemas de partículas con masas, cargas y espines iguales. En el tema de las partículas "idénticas", aparecía una limitación de la individualidad implicada en los conceptos corpusculares clásicos.

Unos meses antes de la conferencia, Heisenberg había formulado el llamado principio de incertidumbre, significativa contribución a la elucidación del contenido físico de la mecánica cuántica. El lector hallará en el primer artículo de esta serie sobre la revolución cuántica (INVESTIGACION Y CIENCIA, julio) una detallada explicación de dicho principio, punto clave de la reunión de Bruselas.

Para introducir la discusión, Bohr presentó un resumen sobre los problemas epistemológicos que enfrentaban a los físicos y aprovechó esta oportunidad para entrar en la cuestión terminológica. Asimismo, precisó el alcance de la noción de complementariedad, de la que ya había hablado hacía poco tiempo en una conferencia anterior que tuvo lugar en Como, y de la que también puede hallarse una completa reseña en el segundo artículo de la serie (INVESTIGACION Y CIENCIA, octubre).

Desde los comienzos del Congreso, Einstein y su posible actitud respecto de los problemas planteados constituían una verdadera incógnita. Todos los presentes tenían enorme curiosidad por conocer sus puntos de vista que, como vamos a ver, diferían considerablemente de los de Bohr.

No debe olvidarse que Einstein no pudo estar presente en Como; de ahí su vivo interés personal en conocer las nuevas ideas. Las ambigüedades de la terminología supusieron el primer escollo de consideración para llegar a un acuerdo en torno a los problemas epistemológicos planteados. Esta situación la expresó humorísticamente Ehrenfest, quien escribió en la pizarra un fragmento bíblico alusivo a la confusión de lenguas en la Torre de Babel. Los intercambios de opiniones proseguían entre pequeños grupos en los descansos. (Bohr recordó en una ocasión haber tenido charlas muy provechosas con Einstein en los pasillos del Instituto Solvay.)

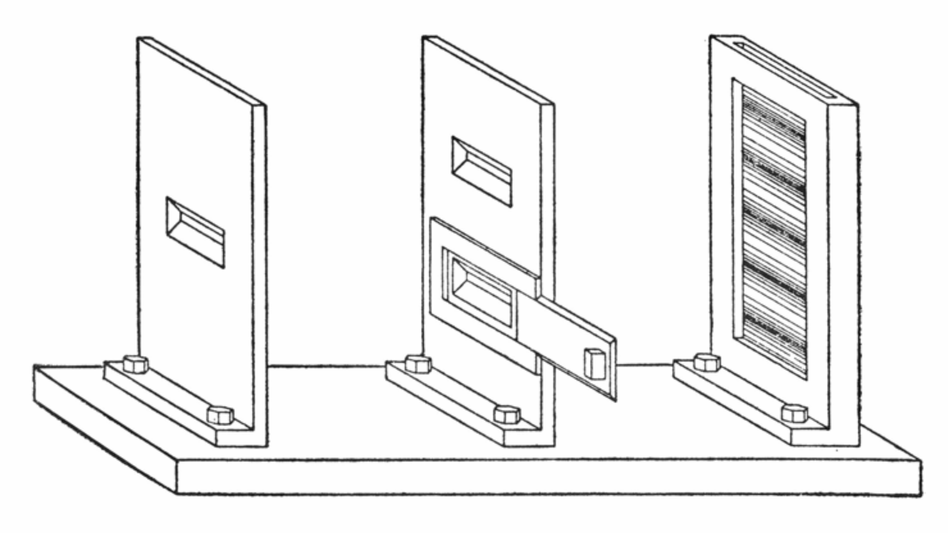
El centro de la cuestión radicaba en saber si la descripción mecánico-cuántica agotaba todas las posibilidades en punto a dar cuenta de los fenómenos observables o si, como mantenía Einstein, podía llevarse a cabo un análisis más profundo que permitiese obtener una descripción más completa de los mismos. Para ello consideraba detalladamente cuál era el balance de la energía y del momento en los procesos individuales, distinguiendo entre el sistema en estudio y los aparatos de medición. Vale la pena analizar un ejemplo sencillo para hacernos una idea de cuáles eran los argumentos esgrimidos en la discusión. Sea el mismo caso que se consideró en la conferencia de Bruselas: una partícula que penetre por un diafragma, es decir, por una rendija, la cual pudiera abrirse y cerrarse a voluntad.

Representemos el estado de movimiento de la partícula antes de alcanzar el diafragma mediante un "tren" de ondas planas. Relacionemos el momento, P, de la partícula con el número de ondas, s, mediante la ecuación $P=h\,s$, obtenida por Einstein en 1905 al estudiar el efecto fotoeléctrico, utilizando los famosos "cuantos de luz". De acuerdo con la difracción que sufren las ondas al atravesar la rendija, representemos el estado de movimiento de la partícula mediante ondas esféricas, con una apertura angular bien definida y con una extensión radial limitada, caso de usar la posibilidad de abrir y cerrar la rendija durante el proceso. Aparece, en el primer caso, una componente del momento de la partícula, Δp, en el plano del diafragma, que se debe a la interacción con el mismo. En el segundo caso, cuando se usa la "ventana", tenemos una energía cinética adicional, ΔE, para nuestra partícula. En el plano del diafragma, registramos también una indeterminación en la localización de la partícula, que viene medida por la anchura de la rendija. Con esto queremos decir que cometemos un error igual al valor de dicha anchura al afirmar, por ejemplo, que la partícula pasa por el centro de la rendija. Podemos obtener ahora una relación entre Δp y

esta última magnitud, usando la igualdad que relaciona la apertura angular, la anchura de la rendija y el número de ondas. Quedaría exactamente el principio de incertidumbre de Heisenberg: $\Delta p \approx h/\Delta q$. Para el segundo caso, y por procedimientos análogos, llegaríamos a una relación semejante entre las variables conjugadas E y t: $\Delta E.\ \Delta t \approx h$. Medimos aquí el tiempo t existente entre la apertura y el cierre de la rendija. Einstein sugirió entonces la siguiente cuestión: ¿podía controlarse, y por tanto conocerse mejor, el estado de la partícula después de atravesar la rendija, la transferencia de momento y energía entre la partícula y el diafragma (amén del esfuerzo por localizar a la partícula en el espacio y en el tiempo)? Según el principio de incertidumbre, que Einstein no acababa de aceptar, a pesar de que se lo habían demostrado utilizando sus propias fórmulas, un control preciso del intercambio de momento y energía entre la partícula y los dispositivos utilizados supone perder la posibilidad de localizarlos con igual precisión en el espacio y en el tiempo.

Se propuso entonces, durante esa misma discusión, introducir un segundo diafragma, a continuación del primero, con dos rendijas en lugar de una. En condiciones normales, obtendríamos, en una placa fotográfica destinada al efecto, una figura de interferencias; fenómeno éste provocado por los caminos que pueden seguir en este caso las ondas. Por otra parte, con este dispositivo, nuestra partícula tiene también dos posibilidades de recorrer el trayecto hasta la placa. Según la transferencia de momento que tenga lugar entre la partícula y el primer diafragma, ésta seguirá uno u otro cambio, el superior o el inferior. Einstein opinaba que controlando dicha transferencia se permitiría un mejor análisis del fenómeno y, en particular, conocer el itinerario seguido por la partícula. Examinando la cuestión con mayor detalle, puede verse sin embargo que dicho control sugerido por Einstein supone una indeterminación en la posición del diafragma que excluye la aparición del fenómeno de interferencia. La argumentación puede llevarse a cabo, como lo hizo Bohr, usando de nuevo las ecuaciones que obtuvo Einstein en 1905 y ciertos resultados de teoría de ondas.

Las observaciones de Einstein instigaron el ánimo de Bohr, quien diseñó años más tarde una serie de dispositivos "seudorrealistas" según sus propias palabras; en ellos pretendió ilustrar el carácter mutuamente exclusivo de las condiciones bajo las cuales aparecen los fenómenos complementarios. Uno de ellos puede verse en la ilustración adjunta; está relacionado directamente con el experimento antes descrito. La posición relativa de los difragmas y de la placa fotográfica está asegurada por una conexión rígida a base de tuercas. Obviamente, resulta imposible controlar así el intercambio de momento entre la partícula y las diferentes partes del aparato. Para ello necesitaríamos eliminar las tuercas que aseguran los diafragmas, con lo cual la posición de los mismos quedaría indeterminada. La única manera que tenemos de hacer pasar la partícula por una de las dos rendijas del segundo diafragma estriba en cerrar la otra rendija con una placa. En este caso, no se producirían fenómenos de interferencia, ya que las ondas sólo podrían seguir un camino, y no dos como ocurría antes. Observaríamos entonces una distribución continua, semejante a la obtenida con un solo difragma. Estos aparatos de carácter instructivo, ideados por Bohr, señalan la distinción entre los objetos que investigamos y los instrumentos de medida que sirven para definir, en términos clásicos, las condiciones bajo las cuales aparecen los fenómenos. Para Bohr, como él mismo indicó repetidas veces, el fenómeno depende en cualquier caso de la totalidad del dispositivo experimental: la introducción de cualquier pieza en el aparato, un espejo por ejemplo, alteraría sustancialmente las predicciones sobre los resultados que se esperan obtener.



Aparato "seudorrealista" de Niels Bohr

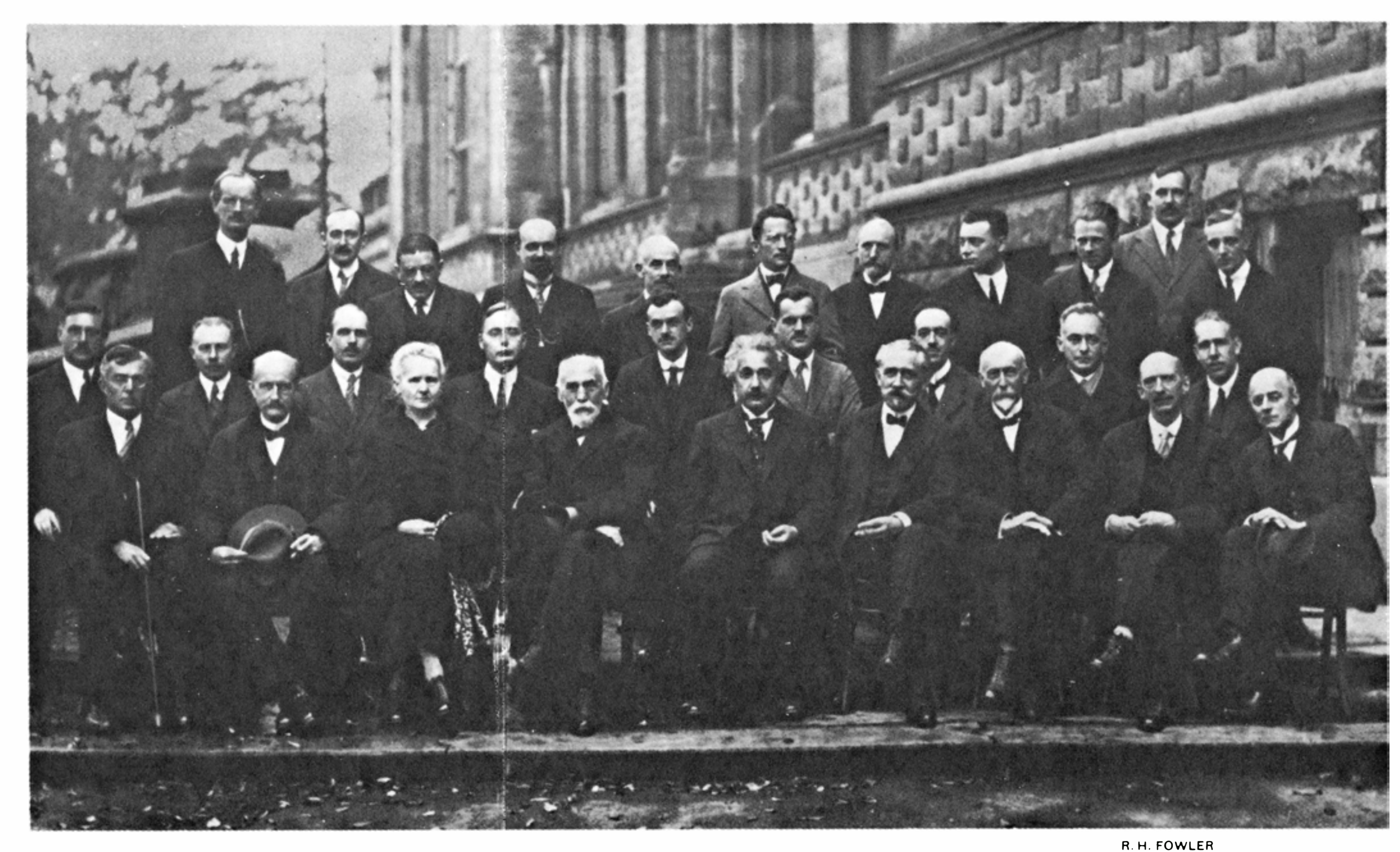
Valga otro ejemplo, que interesó vivamente a Einstein. En él se muestra cómo la imposibilidad de subdividir los fenómenos atómicos impone renunciar a la visualización de los mismos. Si situamos un espejo semirreflectante en la trayectoria de un fotón, dejándole de este modo dos posibles direcciones de propagación, y colocamos dos placas fotográficas a gran distancia en ambas direcciones, dicho fotón podrá registrarse en una de ellas y solamente en una. Si, por el contrario, utilizamos espejos en lugar de placas, podremos observar efectos de interferencia entre los dos trenes de ondas reflejados. En cualquier intento de representar el comportamiento del fotón, nos encontramos siempre con la misma dificultad: estamos obligados a decir por un lado que el fotón siempre escoge uno de los dos caminos y, por otro lado, que se comporta como si hubiera pasado por los dos.

En la siguiente Conferencia Solvay de 1930, la dicusión entre Einstein y Bohr alcanzó tintes dramáticos, como señala el primero en uno de sus libros. En esta ocasión, Einstein recurrió a su famosa fórmula que relaciona masa y energía: $E = m\,c^2$. Se trataba por consiguiente de recurrir a los resultados de la relatividad especial, campo que abrió el propio Einstein en 1905 con un artículo titulado "Sobre la electrodinámica de los cuerpos en movimiento". Mediante una simple pesada, podemos medir la energía total de cualquier sistema y controlar así, en principio, la energía transferida al mismo al interactuar con un objeto atómico. Einstein consideró una caja con una rendija en su interior que podía abrirse y cerrarse mediante un dispositivo conectado a un reloj. Si, inicialmente, preparamos el reloj para abrir el orificio durante un período de tiempo muy corto y en un instante dado, podemos afirmar que un fotón ha atravesado la rendija en un momento que podemos determinar con toda la precisión que deseemos. Más aún, parece posible, midiendo lo que pesa la caja

antes y después de la salida del fotón, calcular simultáneamente la energía del mismo.

Todo ello, evidentemente, está en contradicción con el principio de incertidumbre de la mecánica cuántica. La discusión se centró entonces en un dispositivo del tipo "seudorrealista", ideado por Bohr de acuerdo con las indicaciones de Einstein. El punto crucial radicaba ahora en usar la teoría de la relatividad general, que, como es bien sabido, también la desarrolló Einstein en 1915, para rebatir los argumentos anteriormente expuestos. En primer lugar, Bohr hizo observar las consecuencias que podía tener la identificación entre masa inercial y masa gravitatoria, contenida en las fórmulas de la relatividad especial. Había que tener en cuenta también la relación entre la velocidad de las agujas del reloj y su posición en el campo gravitatorio, deducida del principio de equivalencia de Einstein entre los efectos gravitatorios y los fenómenos observados en los sistemas de referencia acelerados. De este modo, al suspender la caja de un dispositivo adecuado para pesarla con gran precisión, se introduce una indeterminación en la medida del tiempo en que se abre la rendija, al haber variado la posición de la caja en el campo gravitacional y, por consiguiente, la velocidad de las agujas del reloj. La discusión, tan ilustrativa de la fuerza y de la consistencia de los argumentos relativistas, volvía a poner de manifiesto, una vez más, la necesidad de distinguir en el estudio de los fenómenos atómicos entre los propios instrumentos de medida que sirven para definir el sistema de referencia y aquellos elementos que deben ser considerados como objetos de investigación, para los cuales no pueden despreciarse los efectos cuánticos.

Aunque podría pensarse lo contrario, Einstein manifestó, a raíz mismo de la conferencia, no sentirse del todo satisfecho con las explicaciones de Bohr, y señaló nuevos aspectos de la situación que acentuaban su actitud crítica. Debido a la situación política reinante, Einstein no pudo asistir a la Conferencia Solvay de 1933, consagrada a los problemas sobre la estructura y propiedades de los núcleos atómicos; tuvo, sin embargo, ocasión de exponer sus nuevas ideas en una comunicación publicada en 1935. Escrita en colaboración con Podolsky y Rosen, se titulaba: "¿Puede considerarse completa la descripción mecánico-cuántica de la realidad física?". El criterio usado como argumentación en dicho apunte era el siguiente: "Si, dejando inalterado el sistema, podemos predecir con seguridad (es decir con probabilidad igual a la unidad) el valor de una cantidad física, entonces existe un elemento de la realidad física correspondiente a dicha cantidad física". Los autores concluían que la mecánica cuántica no proporcionaba una descripción completa de los fenómenos y que, por consiguiente, había que desarrollar una nueva forma de dar cuenta de la realidad física. La réplica de Bohr no se hizo esperar; en un artículo dado a luz ese mismo año refutó las afirmaciones de Einstein y colaboradores, recurriendo de nuevo al punto de vista de la complementariedad y usando los mismos razonamientos que



R. H. FOWLER
A. PICCARD E. HENRIOT ED. HERZEN TH. DE DONDER E. SCHROEDINGER W. PAULI W. HEISENBERG L. BRILLOUIN
P. EHRENFEST E. VERSCHAFFELT
DEBYE M. KNUDSEN W. L. BRAGG H. A. KRAMERS P. A. M. DIRAC A. H. COMPTON L. V. DE BROGLIE M. BORN N. BOHR
I. LANGMEIR M. PLANCK MADAME CURIE H. A. LORENTZ A. EINSTEIN P. LANGEVIN CH. E. GUYE C. T. R. WILSON
O. W. RICHARDSON
ABSENTS: SIR W. H. BRAGG, MM. H. DESLANDRES ET E. VAN AUBEL

V Consejo Solvay de Física. Bruselas, 24-29 de octubre de 1927

en la reunión de Bruselas. Subrayaba principalmente la ambigüedad del significado de la expresión "dejando inalterado el sistema" y la importancia de las verdaderas condiciones que definen los posibles tipos de predicciones sobre el comportamiento futuro del sistema. Para Bohr, estas condiciones constituyen un elemento inherente a la descripción de cualquier fenómeno que podamos considerar como realidad física propiamente dicha. No se justifica entonces la afirmación de que la descripción mecánico-cuántica sea esencialmente incompleta.

Como centro de toda la discusión aparecen dos cuestiones clave: el problema del conocimiento humano, que preocupaba enormemente a Einstein, y el de la terminología. Nos parece interesante citar aquí una definición dada por Bohr de "fenómeno": "Como modo más apropiado de expresarse, abogo por la aplicación de la palabra fenómeno exclusivamente referida a observaciones obtenidas bajo circunstancias especificadas, incluyendo la totalidad del dispositivo experimental". Se trata, pues, de eliminar la distinción entre los procesos aislados, sujetos a efectos cuánticos, y los aparatos de medición, elaborados de acuerdo con conceptos clásicos.

Aun sin encontrar ninguna contradicción formal en la interpretación de la mecánica cuántica de la Escuela de Copenhague, Einstein, insistiendo en su modo particular de entender la física, afirmaba en 1949: "Lo que no me satisface en esta teoría, desde un principio, es su actitud hacia lo que yo considero como uno de los objetivos programáticos de la física: la descripción completa de cualquier situación (individual) real (como se supone que existe fuera de cualquier observación o substanciación)". Aunque aceptadas ya las nuevas ideas desde el Consejo Solvay de 1927, la postura de Einstein resultó muy provechosa de cara al perfeccionamiento de dichas ideas, y es curioso observar cómo una mentalidad revolucionaria como la suya se negaba a aceptar, aferrándose a ideas clásicas, la revolución cuántica defendida por Bohr, exponente de una nueva forma de concebir la realidad física. El problema sigue abierto y cualquier nueva teoría que pueda surgir deberá tenerlo muy en cuenta. (B. A.)

Prevención de la subnormalidad

La bioquímica de la subnormalidad sienta las bases de cuáles han de ser las medidas preventivas de esa deficiencia generalizada (véase INVESTIGACION Y CIENCIA, julio). Las principales medidas que pueden adoptarse en la actualidad para la prevención de la deficiencia mental son:

1) *Antes de la concepción:* Vacunaciones (Vacunación antirrubéolica) y examen de las infecciones, del grupo sanguíneo y Rh, de las afecciones maternas y de los genotipos anormales.

2) *Durante la gestación:* Control médico periódico del curso del embarazo, evitación de agentes mutagénicos (medicamentos, radiaciones, etc.) e infecciones y bienestar fetal; diagnóstico antenatal, afecciones maternas (toxemia, hipertensión, etc.) y adecuada nutrición.

3) *Durante el parto:* Evitación de traumatismo cerebral e hipoxia, monitorización de los partos con riesgo fetal, elevado, índices de sufrimiento fetal, atención especial a los hipo e hipermaduros, reanimación adecuada del recién nacido y unidades de vigilancia intensiva del recién nacido.

4) *Durante el período neonatal:* Determinación de las metabolopatías; determinación de la hiperbilirrubinemia, prevención de las infecciones y alimentación adecuada.

Los grupos de alto riesgo son aquellos que han tenido antecedentes de subnormalidad. El consejo genético se desarrollará más en la medida en que tener hijos constituya una elección activa sopesada y conscientemente adoptada por los progenitores. Los antecedentes raciales, sus costumbres y medio de vida deben tenerse, asimismo, en cuenta.

El examen inicial a las embarazadas debe establecer, entre otros datos, los siguientes: fecha de comienzo de embarazo, situación socioeconómica de la madre, historia médica de la embarazada, exploración clínica con análisis completo de sangre comprendiendo las pruebas de lúes y el factor Rh, así como grupos sanguíneos, investigación de anticuerpos contra la rubéola, test cutáneo para la detección de toxoplasmosis, detección de enfermedades cardíacas y trastornos endocrinos, bacteriología urinaria, historia ginecológica de la madre y pruebas citológicas. En sucesivos exámenes deberá añadirse a los datos anteriores el control de medicamentos administrados a la embarazada, control de drogas no terapéuticas, tabaco, alcohol y otros.

Las circunstancias socioeconómicas son muy importantes en la consideración del período perinatal y sus posibles secuelas. La frecuencia relativa del retraso mental "límite" se halla claramente relacionada con una situación socioeconómica baja. Las visitas al médico durante

el embarazo son menos (cuando se realiza alguna), se tienen menos en cuenta los antecedentes familiares o de anteriores embarazos, la tensión neurológica es peor.... Asimismo, el nivel cultural de los padres influye sensiblemente en las medidas que se adoptan, en el ambiente en que el niño se desarrolla. La educación sanitaria de la población es, pues, absolutamente imprescindible.

Según Royer, las medidas actuales que permiten el consejo genético, la prevención y tratamiento de algunas alteraciones moleculares congénitas son: la detección de anormalidades en el periodo neonatal, tales como la fenilcetonuria, galactosemia o leucinosis; el diagnóstico prenatal, a través de amniocentesis y examen del líquido amniótico o de las células cultivadas *in vitro;* el reconocimiento y la caracterización de los heterocigóticos. En tanto no pueda realizarse la corrección o substitución del mensaje genético alterado en la especie humana, las posibilidades terapéuticas principales pueden resumirse así:

a) Reemplazamiento de una macromolécula ausente o anormal; este sistema se ha aplicado con éxito en el tratamiento de la hemofilia, afibrinogenemia y anormalidades genéticas de las inmunoglobulinas.

b) Reemplazamiento de un enzima de acción digestiva. Las anomalías hereditarias de biosíntesis de enzimas elaboradas por las células del páncreas exocrino o por la mucosa del intestino tienen dos tratamientos bien claros, a saber: suprimir el alimento precursor cuando puede ser substituido por un equivalente que proporcione los mismos requerimientos nutritivos o bien administrar el enzima deficitario por vía oral (tripsinógeno pancreático en el caso de déficit en lipasa pancreática, o administración de maltasa en los casos de intolerancia genética a la sacarosa o isomaltosa).

c) Bloqueo hereditario de una etapa de un ciclo metabólico. De un lado, existen los efectos tóxicos producidos por el acúmulo de los metabolitos, previos a la reacción interrumpida, que pueden inhibir las actividades de otros sistemas enzimáticos, lesionar tejidos, alterar el medio extracelular y otros. Por otra parte, se derivan situaciones carenciales, ya que no se produce la biosíntesis de determinadas sustancias. En consecuencia, en estos casos los tratamientos pueden ser principalmente de tres tipos. En primer lugar, administración de las sustancias cuya biosíntesis ha sido interrumpida por la deficiencia enzimática. Pueden citarse como ejemplo las anomalías hereditarias en la biosíntesis de hormonas corticosuprarrenales, que pueden controlarse eficazmente mediante la administración de corticosteroides y, asimismo, las alteraciones hereditarias en la síntesis de hormonas tiroideas, que pueden ser compensadas mediante la administración de tiroxina. En segundo lugar, la supresión de un nutriente que actúa como sustrato antes de la reacción bloqueada. El ejemplo más famoso es el de la fenilcetonuria, en el cual se aplica una dieta pobre en fenilalanina. Por último, la creación de un bloqueo metabólico supletorio, con el fin de evitar la formación de una sustancia nociva: en el caso de la enfermedad de Lesch-Nyhan, la administración de alopurinol inhibe la acción de la xantina oxidasa, que permite suprimir el riesgo litógeno de la hiperxantinuria, mucho más grave que el de la hiperuricemia. Sin embargo, este tratamiento no evita la encefalopatía grave que conlleva la enfermedad.

d) Anomalías en las hidrolasas lisosómicas. En estos momentos no existe solución terapéutica para este tipo de anomalías. Como ensayos en este sentido, deben mencionarse las transfusiones de plasma en el caso de la enfermedad de Fabry y en la mucopolisacaridosis.

e) Anomalías en los sistemas de transporte transcelular. Estas anomalías hereditarias afectan principalmente al intestino, al túbulo renal o ambos. Son diversas las enfermedades que pueden catalogarse dentro de este grupo: cistinuria, hipercistinuria, hiperaminoaciduria, enfermedad de Hartnup, iminoglicinuria, glicinuria, etcétera. Estas enfermedades, cuyas bases moleculares son todavía precariamente conocidas en algunos casos, se tratan actualmente mediante modificaciones dietéticas.

f) En el caso de alteraciones moleculares en las proteínas hormonorreceptoras, los tratamientos aplicados son también diversos. Las enfermedades de esta naturaleza mejor conocidas son los defectos hereditarios de recepción de señales hormonales por células tubulares del riñón. En el plano de la expresión sintomática son seudoendocrinopatías. Una de ellas es la insensibilidad tubular a la vasopresina, o diabetes insípida nefrogénica, que se transmite recesivamente ligada al sexo. Esta alteración se mejora por un aporte importante y permanente de agua y un régimen pobre en componentes osmóticamente activos; si bien debe señalarse que algunas sustancias y, en especial la hidroclorodiazida, simulan la acción de la hormona, disminuyendo,

como lo hace la vasopresina, la eliminación de agua. La insensibilidad tubular a la paratohormona, enfermedad de transmisión dominante ligada al cromosoma X, produce como fenotipo el seudohipoparatiroidismo. La paratohormona activa el enzima adenilciclasa de la corteza renal. Esta enfermedad se mejora sustancialmente mediante la administración de calciferol en altas dosis, que permite rebajar el porcentaje de reabsorción tubular del fósforo, hasta el punto de simular el efecto de la hormona.

La terapéutica de reemplazamiento, es decir, la de cambiar lo que no funciona por una molécula normal, puede realizarse únicamente, de manera directa, cuando la alteración se centra en el torrente circulatorio. Ejemplos de esta situación son la hemofilia, agammaglobulinemia y analbuminemia. Sin embargo, si el enzima inhábil se halla en algún orgánulo intracelular, donde ejerce su función específica, no es posible su reemplazamiento por aplicación directa, ya que su administración al torrente circulatorio no redundaría en su entrada en la célula, en la cual ella ejerce su acción.

Otra posibilidad consiste en la inducción enzimática: así, por ejemplo, el fenobarbital induce la glucuroniltransferasa hepática; de ahí su aplicación para disminuir la concentración plasmática de bilirrubina no conjugada en algunos recién nacidos afectos de hiperbilirrubinemia hereditaria, que cursa con ictericia.

Citemos también los tratamientos con vitaminas. Cuando la anomalía molecular afecta a la unión con el grupo prostético coenzimático, o cuando la enzimopatía se refiere a alguna de las etapas propias de la biosíntesis del coenzima activo, la administración de dichos componentes coenzimáticos constituye un magnífico tratamiento terapéutico. Pueden citarse como ejemplos la administración de tiamina (vitamina B) en el caso de la acidosis láctica crónica hereditaria con hiper-α-alaninemia; la administración de biotina en determinadas acidemias propiónicas; la administración de vitamina D en la forma hipocalcémica de raquitismo, de transmisión recesiva, y otros. Sin embargo, los dos ejemplos más interesantes se refieren a la vitamina B_6 y a la vitamina B_{12}. Se conocen alrededor de media docena de enfermedades hereditarias que dependen de la piridoxina, tales: las convulsiones neonatales dependientes de piridoxina, donde existe un déficit congénito en descarboxilasa del ácido glutámico; la anemia familar microcitaria por déficit de la δ-aminolevunílico sintetasa; la cistationuria por déficit en cistationasa; algunas formas de homocistinuria por déficit de cistationina sintetasa, etcétera. También son complejas las reacciones en las que interviene la vitamina B_{12}. La metilcobalamina y la 5'-desoxiadenosilcobalamina actúan de coenzimas en dos reacciones: la remetilación de la homocisteína en metionina y la isomerización del metilmalonil coenzima A en succinil-coenzima A.

La terapéutica con microcápsulas es esperanzadora. El enzima deficitario se "enjaula" en una estructura microscópica de polímero, de modo que, a través de sus poros, puedan acceder fácilmente al interior los sustratos y cofactores de pequeño tamaño, pero no las proteínas y otras moléculas de peso molecular elevado, evitándose por tanto su capacidad antigénica. Se han "encapsulado" enzimas tales como la ureasa, asparaginasa y catalasa. La ureasa se emplearía para corregir la enzimopatía correspondiente, pero también para actuar en el riñón artificial; la asparaginasa para impedir el crecimiento de algunos tumores dependientes de L-asparagina; y la cetolasa para tratar los casos de acatalasemia, que es la falta o disminución del enzima catalasa en sangre.

Se preparan microesferas recubiertas de una membrana de nylon, con poros de 16 Å de diámetro. También pueden prepararse "liposomas". La eficiencia de las microcápsulas en la hemodiálisis es evidente, si tenemos en cuenta que microcápsulas de un diametro medio de 20 micras en un volumen total de 10 ml poseen una superficie de 25.000 cm^2 (!), es decir, muy superior a un riñón artificial convencional, que ocupa un lugar muchísimo mayor.

La incidencia de las enzimopatías evitables (fenilcetonuria, homocistinurias y otras) es ciertamente escasa, si se compara con las lesiones cerebrales producidas por otras causas (cromosomopatías, trauma de parto, etcétera). Sin embargo, una adecuada atención a las características bioquímicas durante las principales etapas del desarrollo podría conducir a evitar las alteraciones patológicas que se producen como consecuencia de una clara predisposición hereditaria, que pone de manifiesto a su vez un repertorio enzimático característico. Tal es el caso de las hiperlipemias, cuya determinación podría aplicar selectivamente los tratamientos dietéticos adecuados, con el fin de evitar un porcentaje considerable de ulteriores accidentes cardiovasculares.

Los programas de selección genética, ha escrito Scriver, tienen por objetivo identificar aquellos individuos a quienes sus alelos particulares sitúan (a ellos o a sus descendientes) en riesgo de adquirir determinadas enfermedades, que no se presentan –por carecer de esta singularidad genética– en sus vecinos o familiares. ¿Cómo puede el conocimiento más profundo de la genética resolver estos conflictos entre el individuo y la sociedad? ¿Puede incrementar dicho conocimiento la eficacia de la medicina preventiva? Ciertamente, la genética puede resolver estos problemas al poner de manifiesto la individualidad biológica de los ciudadanos. No es necesario hacerlo en todos los hombres, sino en aquellos de mayor riesgo.

El riesgo promedio de una población en relación a una alteración determinada viene dado, a la postre, por el conjunto de quienes tienen mucho riesgo y quienes tienen muy poco. Las medidas preventivas deben dirigirse, evidentemente, a las personas de mayor riesgo. La educación de una medida preventiva –la administración de un medicamento, el seguimiento de una dieta u otros– será mucho mejor cuanto más específico, ya que la adopción de medidas generales, dirigidas uniformemente a todos los ciudadanos, resultarían en muchas ocasiones, aparte de innecesarias, perjudiciales.

La influencia del medio cultural es extraordinaria, no debiendo restringir, en consecuencia, la consideración del problema a las medidas que pueden adoptarse en cada uno de los tipos de prevención mencionados, si no a la propia asequibilidad de conocimientos, al suministro correcto de la información que permitan el desarrollo cabal de la personalidad, del potencial hereditario (niños privados de estímulos, faltos de afectividad y defectos de atención e higiene).

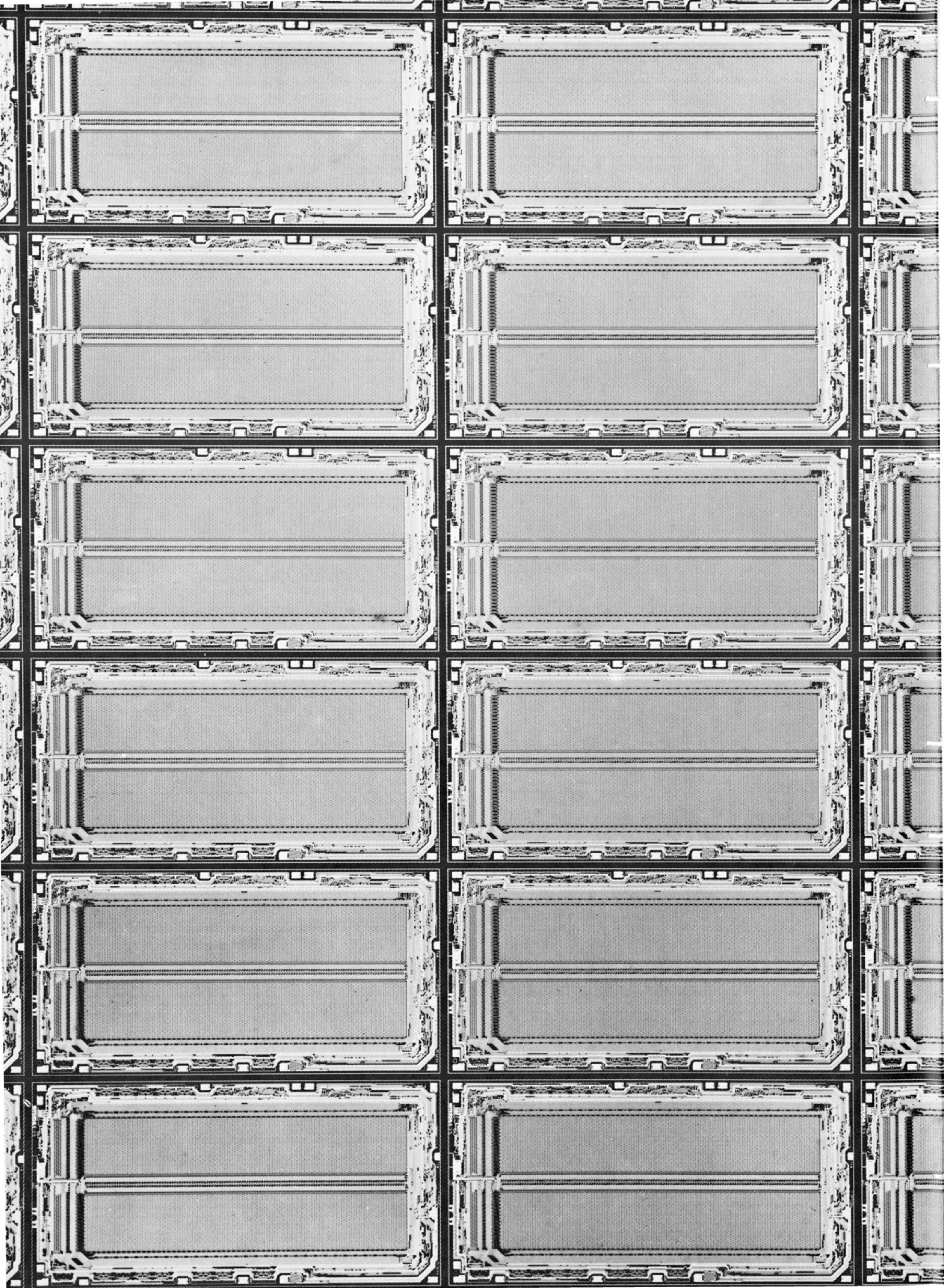
Según datos de Estados Unidos, alrededor del 25 por ciento de los retrasados mentales poseen un cociente intelectual inferior al mínimo requerido para ser atendido en un centro de educación especial.

Si se atendiera debidamente a los niños que requieren educación "vigilada" pero que pueden –y quizás deben– seguir sus estudios con los niños "normales"; si los centros de educación especial tuvieran plaza para todos los que pueden necesitarla... y sólo para ellos (es decir, no deben seguir estudios ni capacitación profesional en estos centros los niños profundos irrecuperables, con tanta frecuencia mezclados, por falta de recursos personales o materiales, con los otros, ni deben seguirlos los que han alcanzado ya un grado de recuperación suficiente para integrarse en niveles "no especiales" de educación); si no existieran barreras arquitectónicas; si se prestara la debida atención social a los mudos, a los sordos, a los ciegos, a los inválidos... a los subnormales, nos daríamos cuenta –entonces, cuando los niños minusválidos tuvieran, al menos, los mismos derechos que los normales– de la extraordinaria rentabilidad de las medidas que pueden adoptarse con los conocimientos ya existentes y de la necesidad de profundizar más en estos conocimientos en lugar de seguir favoreciendo el disfrute de bienes superfluos en la población sana.

A medida que el conocimiento de la variación genética sea mayor, podremos pasar desde una medicina preventiva convencional a una medicina preventiva que tenga en cuenta las características genéticas desde el principio. Si cada hombre es distinto y dotado de capacidades diferentes, el auténtico progreso radicará en tratarle, en cuidarle según estas capacidades, es decir, de acuerdo con su individualidad. La evolución no ha conducido sólo a una especie o a una raza. La evolución es dinámica y opera continuamente en el conjunto y en cada una de sus partes. De tal modo que cada uno no es siquiera idéntico –dado el inmenso número de mutaciones que sufre su dotación humana– a como era en el momento precedente. Sólo esta compensación de la evolución a nivel individual puede conducir al correcto planteamiento de la personalidad bioquímica y, sobre todo, por ser la que más nos atañe, de la personalidad humana.

Poco a poco, se irá "personalizando" la atención médica gracias a los datos bioquímicos y citológicos que permitan configurar el perfil clínico de cada uno. Se hará una medicina "a la medida".

En España, en 1977, Año internacional del niño, acaba de ser aprobado un Plan Nacional de Prevención de la Subnormalidad, realizado bajo los auspicios de la reina Sofía por un numeroso grupo multidisciplinar (obstetras y ginecólogos, neuropsiquiatras, pediatras, nutriólogos, bioquímicos, neonatólogos, etc). Es un buen augurio, porque recuerda la "Declaración de los Derechos del Niño", recogida explícitamente en la Declaración Universal de Derechos del Hombre: "La maternidad y la infancia tienen derecho a una ayuda y una asistencia especiales"... La sociedad tiene, pues, el deber de atenderlos. (F. M.)



La fabricación de microcircuitos electrónicos

Los patrones de los circuitos integrados se hacen primero a gran escala y luego se graban fotográficamente en la pastilla. El objetivo de los métodos de fabricación es la producción a menos coste por función

William G. Oldham

La fabricación de circuitos integrados a gran escala (LSI, del inglés "large-scale integrated") tiene como meta fundamental el menor coste posible por función electrónica desempeñada. Las principales características de los procesos adoptados por la industria microelectrónica se pueden comprender mejor en función de ese objetivo. Entre tales rasgos se cuentan la fabricación de muchos circuitos al mismo tiempo (un ejemplo extraordinario de producción en masa), la reducción de los circuitos al menor tamaño posible y la máxima simplificación de la tecnología del proceso.

La drástica reducción del coste de los microcircuitos electrónicos conseguida en los últimos años no se debe a ningún avance especial en la tecnología de fabricación. Es más, la mayoría de los procesos fundamentales de fabricación son de aplicación generalizada en esta industria desde hace cinco o más años. La drástica disminución reciente en los costes de fabricación se ha conseguido en un período de inflación económica generalizada. El coste de producción de una "oblea" de silicio, que es el substrato en el que se realizan los microcircuitos electrónicos, ha crecido moderadamente, pero el área de las obleas ha aumentado más rápidamente, casi doblándose cada cuatro años. De esta forma, el coste de producción por unidad de área ha disminuido. Mientras tanto, el espacio requerido por una función electrónica dada ha menguado en un factor de dos cada 18 meses, más o menos. Esta reducción de tamaño obedece a una progresiva miniaturización de los elementos del circuito y sus interconexiones, además de suponer un gran ingenio en el diseño de circuitos sencillos y procesos tecnológicos simples para enmascararlos. Es más, la gradual eliminación de defectos en las distintas fases de fabricación ha desembocado en una disminución importante del coste de fabricación neto. Con una frecuencia de defectos menor, la producción de buenos circuitos en una oblea dada aumenta.

MASCARA FOTOGRAFICA. Placa de vidrio que contiene el patrón de una sola capa de circuitería integrada, grabada sobre su superficie en una película de cromo. En estas dos fotografías se muestra a contraluz; una reproduce su tamaño actual *(arriba)* y la otra aumentado 10 veces *(página opuesta)*. Se utiliza en la confección de una RAM (memoria de acceso aleatorio) de 16.384 bits.

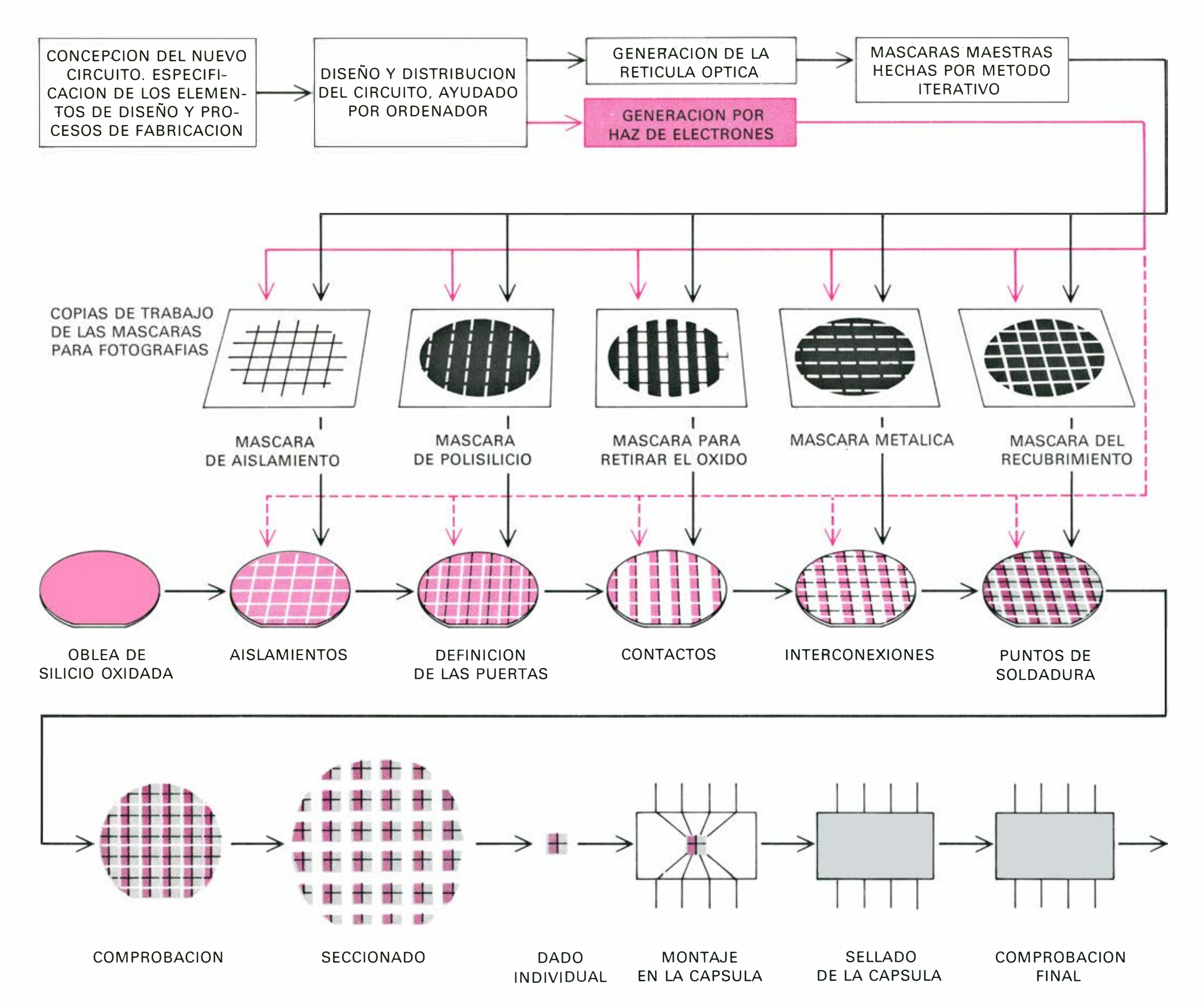
Las tendencias actuales de las investigaciones en la fabricación de microcircuitos electrónicos sugieren que los progresos en la reducción de los costes de producción van a continuar. Virtualmente cada etapa de fabricación –de la fotolitografía al encapsulado– está en medio o en las puertas de un avance significativo.

Un circuito integrado a gran escala contiene decenas de millares de elementos; cada uno de éstos resulta tan exiguo

que el circuito completo no suele superar a una figura geométrica cuadrada de 6 mm de lado. Las obleas de monocristal de silicio puro, que soportan los circuitos, son mucho mayores; miden unos ocho o diez centímetros de diámetro. Una de las claves de la economía de la fabricación de microcircuitos electrónicos es la producción simultánea de cientos de circuitos, uno junto a otro, en una sola oblea. Se está consiguiendo una producción en masa todavía mayor en varias etapas de la fabricación, obteniéndose hasta cien obleas juntas en una serie. De esta forma, el coste de la mano de obra y de equipos se reparte entre miles de circuitos, lo cual permite a su vez un coste extremadamente bajo por circuito, que es lo que caracteriza a la microelectrónica.

Cuando una oblea ya ha pasado por la etapa de fabricación queda seccionada en dados individuales, o pastillas; cada una de éstas es un circuito microelectrónico completo. No todos los circuitos funcionarán. No se pueden evitar obleas defectuosas. Un solo defecto basta para arruinar el circuito entero. Por ejemplo, un leve arañazo de milésimas de milímetro de largo puede romper una conexión eléctrica. Física y económicamente, se hace imposible reparar los circuitos defectuosos, por lo que simplemente se desechan.

Tal como cabe suponer, cuanto mayor sea el dado mayor será la probabilidad de que aparezca un defecto que estropee el circuito. El rendimiento (número de circuitos buenos por oblea) disminuye con



LA CONFECCION de un circuito integrado a gran escala (LSI) está detallada en esta figura. El diseño del circuito suele llevarse a cabo con ayuda de un ordenador, que permite determinar la distribución más compacta de los elementos del circuito. La distribución resultante se emplea para preparar el conjunto de máscaras fotográficas, cada una de las cuales contiene el patrón de una sola capa. Normalmente, esto se hace obteniendo primero una ampliación diez veces mayor de cada capa, llamada retícula; ésta se verifica, corrige y regenera hasta su perfección. Una reducción fotográfica de la retícula se reproduce cientos de veces en un proceso recurrente para obtener un conjunto de máscaras maestras al tamaño definitivo, de las cuales se consigue un gran número de placas de trabajo. (La aproximación a la generación del patrón mediante haz de electrones, representada aquí en color, está empezando a reemplazar al método óptico de producción de máscaras; de esta forma se torna posible eliminar dos pasos fotográficos y se graba el patrón directamente en la placa de trabajo, a partir de la información almacenada en la memoria del ordenador. Se espera que, con el tiempo, esta litografía de haz de electrones se introduzca directamente en la fabricación de los propios circuitos.) Las "obleas" de silicio, que sirven de substrato a los circuitos, se obtienen cortando un largo cristal de silicio en finas rebanadas, que se pulen, limpian y oxidan como preparación para el primer paso de la confección. (A lo largo de la longitud del cristal se produce un "chaflán", de esta forma cada oblea tendrá un flanco de referencia paralelo a un plano natural del cristal.) Siguiendo el proceso de fabricación con cinco máscaras, expuesto en este ejemplo muy simplificado, se somete a prueba la oblea para determinar los circuitos buenos; los circuitos defectuosos se marcan con una gota de tinta. De la oblea seccionada se seleccionan los "dados" que contienen circuitos buenos; luego, éstos se montan en cápsulas. Para asegurarse de que el circuito encapsulado funcione correctamente, se realiza todavía una comprobación final.

el tamaño del dado, porque hay menos espacio libre en la oblea para dados grandes y porque los circuitos mayores tienen mayor probabilidad de acarrear un defecto.

En principio, debe parecer más económico construir circuitos muy simples, y por tanto muy pequeños, puesto que la mayoría saldrán buenos. Los circuitos sencillos son baratos; los circuitos lógicos sencillos pueden obtenerse por 10 centavos de dólar cada uno. Sin embargo, también hay que tener en cuenta los costes de comprobación, encapsulado y montaje de los circuitos completos en un sistema electrónico. Una vez se separan los circuitos, rompiendo las obleas en dados, hay que manejar dado a dado. A partir de ese punto, ningún coste de producción (por ejemplo, el de encapsulado o comprobación) se repartirá entre cientos o miles de circuitos. Para circuitos integrados a media escala (MSI, del inglés "medium-scale integrated"), los costes de encapsulado y comprobación dominan sobre los otros costes de producción.

Un sistema microelectrónico típico contiene circuitos de gran y mediana escala de integración. Los costes de diseño y construcción del sistema crecen rápidamente con el número de circuitos empleados. Para rebajar al mínimo el coste total del sistema deberían emplearse, en teoría, un pequeño número de circuitos muy potentes (requiriéndose que cada circuito sea grande) o bien un gran número de circuitos muy baratos (requiriéndose que cada circuito sea pequeño). En muchos sistemas complejos, el coste total mínimo se consigue con los circuitos más caros y mayores, que valen unos 10 dólares cada uno. Si se usasen circuitos menos potentes y más baratos, serían necesarios muchos más para construir el sistema, y los precios de comprobación y montaje tenderían a subir. Por otra parte, si se almacenan en un circuito demasiadas funciones electrónicas, el gran tamaño del dado tendría como consecuencia una producción tan baja por oblea que el coste de cada circuito resultaría prohibitivo.

Tomando 10 dólares como precio de venta típico de un circuito integrado a gran escala, se puede deshacer el camino y encontrar el tamaño óptimo de los dados. Por supuesto, el precio de venta del circuito debe ser suficientemente alto para cubrir no sólo los costes directos de fabricación, sino además los costes de la investigación y desarrollo, ventas y gastos generales; por tanto, es razonable suponer que el coste directo de fabricación estará bastante por debajo, digamos



EN LA DISTRIBUCION DEL MICROCIRCUITO interviene la ayuda de un ordenador. La porción del circuito integrado a gran escala, cuya distribución se está llevando a cabo en esta demostración, es estudiada por el diseñador en la pantalla de un tubo de rayos catódicos. Escribiendo instrucciones en un teclado o redibujando con un "lápiz luminoso", se pueden añadir, quitar e incluso desplazar los elementos del circuito. El terminal aquí mostrado está fabricado por Applicon Inc.



VERIFICACION DE LA DISTRIBUCION inspeccionando una reproducción de la máscara 500 veces mayor que el circuito real. Para cumplir este objetivo se suele generar una retícula, que represente una ampliación diez veces mayor del patrón de una sola capa del dispositivo, en una placa fotográfica de vidrio de 32 centímetros cuadrados. En este caso, una reproducción de la retícula, ampliada en otras 50 veces más se somete a comprobación para verificar que cada elemento del circuito esté correctamente situado y posea el tamaño adecuado. Para preparar las máscaras maestras hay que proyectar una imagen reducida de la retícula en otra placa fotográfica, mediante un sistema de lentes cuya resolución es sumamente alta. El circuito que se contempla aquí en la mesa luminosa es el mismo que se muestra en las dos primeras fotografías de este artículo.



EL CENTRO DE FABRICACION DE OBLEAS debe mantenerse extremadamente limpio, para evitar la contaminación del entorno de fabricación con la más pequeña partícula de polvo. Los trabajadores visten trajes especiales y el aire se filtra y recircula continuamente para mantener al mínimo el nivel de suciedad. Esta fotografía de una "habitación limpia" ha sido tomada en una planta propiedad de Zilog Inc.

5 dólares. Sólo la manipulación de una oblea de silicio cuesta aproximadamente 100 dólares, sin considerar el tamaño de los dados; cuando incluimos los costes de encapsulado y comprobación, los costes totales de fabricación llegan a doblarse. Por esta razón, si se supone que la fabricación de un determinado circuito integrado ha de costar 5 dólares, el tamaño óptimo será el que produzca aproximadamente 40 circuitos buenos por oblea. Actualmente se puede conseguir este rendimiento con dados de unos cinco milímetros de lado. Conviene recalcar que en este ejemplo el porcentaje de circuitos buenos es bastante bajo. El rendimiento es sólo de 40 circuitos buenos entre los 250 que pueden fabricarse en una oblea de 100 milímetros de diámetro.

La estructura de un circuito integrado es compleja, tanto en la topografía de su superficie como en su composición interna. Cada elemento de un circuito de este tipo tiene una intrincada arquitectura tridimensional, que debe reproducirse exactamente en cada circuito. La estructura se compone de varias capas, cada una de las cuales es un patrón detallado. Algunas capas yacen dentro de la oblea de silicio, y otras están apiladas encima. El proceso de fabricación consiste en formar esta secuencia de capas en precisa concordancia con el plan del diseñador del circuito.

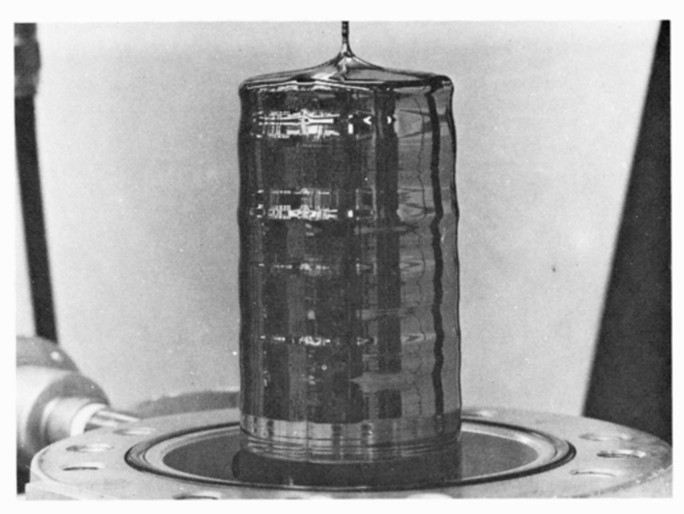
Antes de examinar cómo se forman estas capas, será interesante echar un vistazo general al procedimiento por el cual un circuito integrado pasa de ser una idea del diseñador del circuito a una realidad tangible. En la primera etapa del desarrollo de un nuevo circuito electrónico, los diseñadores que conciben el nuevo producto trabajan en la especificación de las características funcionales del dispositivo. También seleccionan los distintos procesos que se requerirán para fabricarlo. En la siguiente etapa, comienza el diseño real del dispositivo: se calcula el tamaño y la posición aproximada de cada elemento del circuito. La mayor parte de este diseño previo se realiza con ayuda del ordenador.

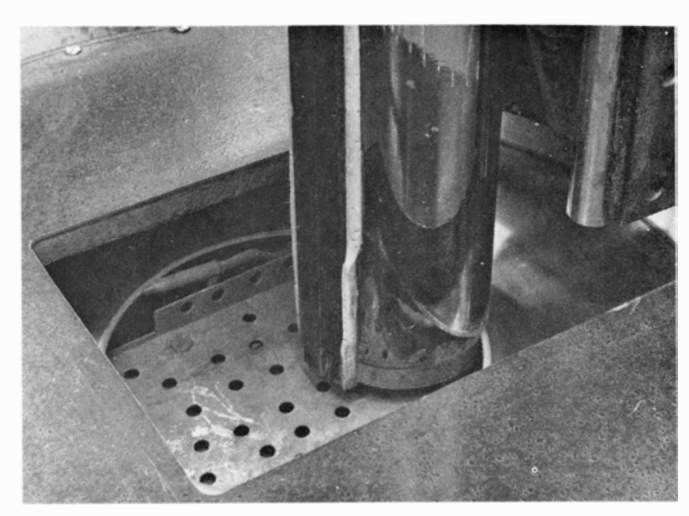
Un ordenador puede simular el funcionamiento de un circuito, de la misma manera que los juegos electrónicos de televisión simulan el desarrollo de una partida de tenis de mesa o de una guerra espacial. El diseñador del circuito supervisa el comportamiento de las tensiones del circuito y ajusta los elementos del circuito hasta que se consigue el comportamiento deseado. La simulación en ordenador resulta más económica que montar y comprobar un prototipo del circuito construido con elementos discretos; además, es más fiable. La ventaja principal de la simulación reside, sin embargo, en que el diseñador puede cambiar un elemento del circuito simplemente escribiendo una corrección en un teclado, e, inmediatamente, comprobar el efecto de la modificación en el comportamiento del circuito.

La distribución final, que da las posiciones precisas de los distintos elementos del circuito, también se realiza con ayuda del ordenador. El diseñador de la distribución trabaja frente a un terminal de ordenador, colocando y desplazando los elementos del circuito, mientras observa

la distribución, ampliada cientos de veces, en la pantalla de un tubo de rayos catódicos. La distribución especifica el patrón de cada capa del circuito integrado. El objetivo de la distribución es conseguir la función deseada de cada circuito en el menor espacio posible. El viejo método de dibujar a mano las distribuciones de los circuitos no ha sido completamente reemplazado por el ordenador. Son muchas las partes de un circuito integrado a gran escala que todavía se dibujan a mano antes de someterlas al ordenador.

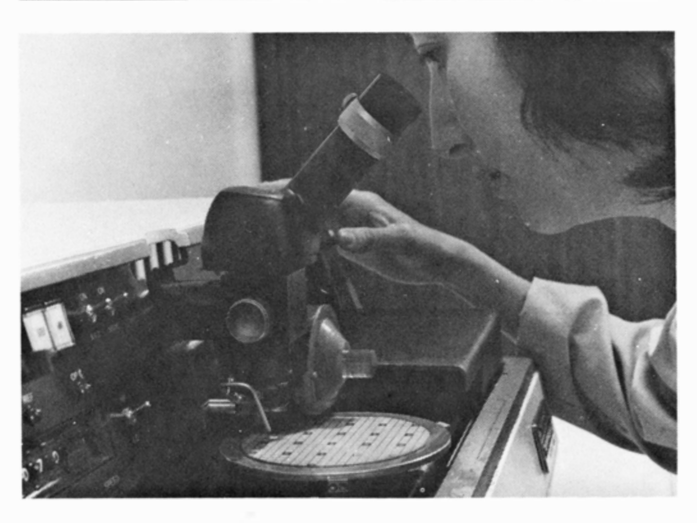
A lo largo de cada etapa de este proceso, incluida la etapa final en que se cierra el circuito, se verifica la distribución mediante detallados diagramas que dibuja el ordenador. Puesto que cada elemento del circuito mide sólo unas micras cuadradas, los diagramas de verificación

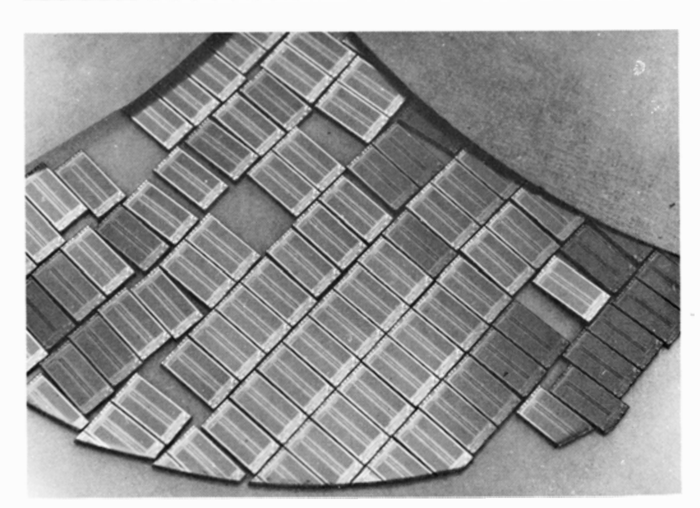




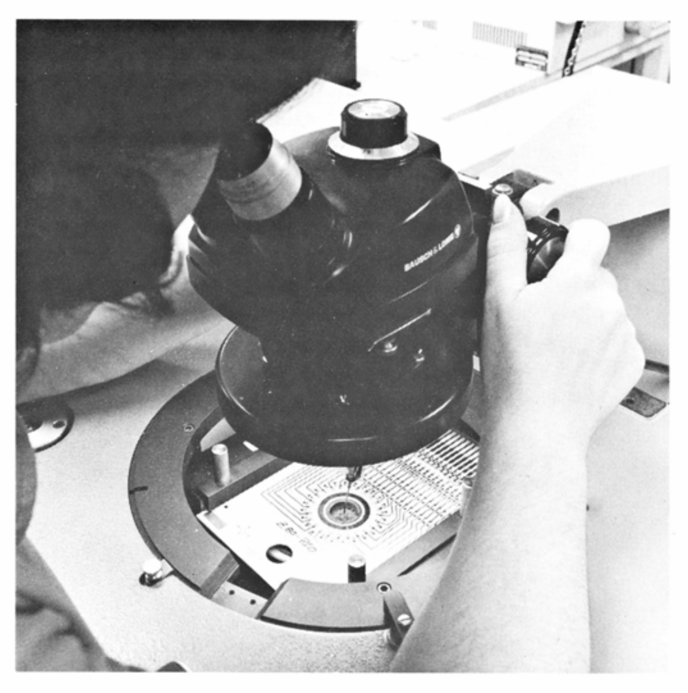


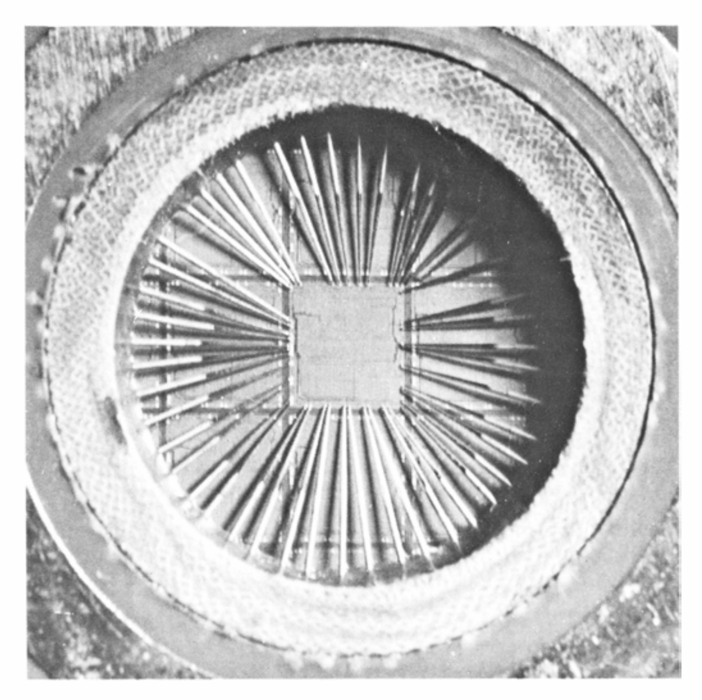




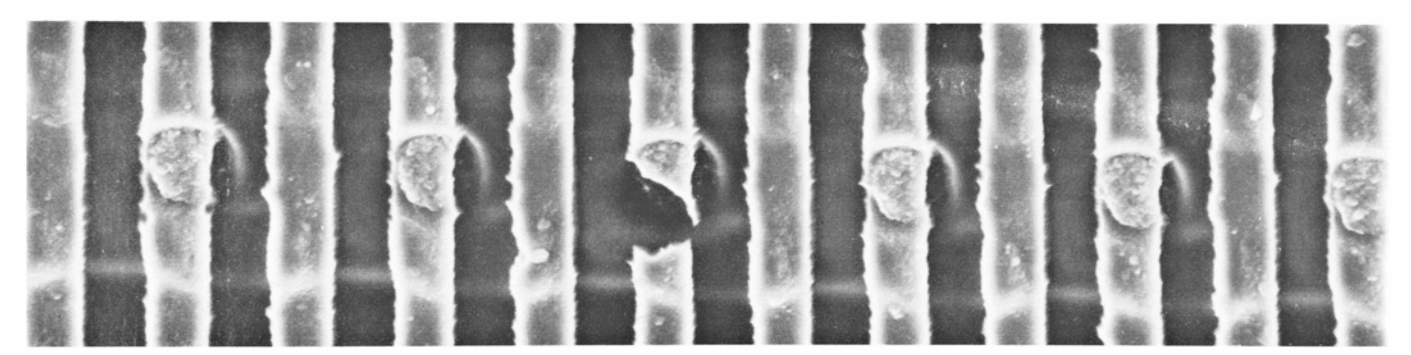


SEIS ETAPAS de la fabricación de un microcircuito electrónico, seriadas en esta secuencia fotográfica: crecimiento de un monocristal de silicio contaminado, retirando una pequeña "semilla" de cristal de silicio a partir de un crisol de cuarzo que contiene silicio puro fundido, al cual se ha añadido el contaminante deseado *(izquierda, arriba);* **rebanado del cristal de silicio, mecanizado con precisión en finas obleas con una sierra circular con filo de diamante** *(derecha, arriba);* **comprobación de la regularidad de la superficie pulida de una oblea en un dispositivo de medida de precisión** *(izquierda, centro);* **formación de una delgada capa superficial de dióxido de silicio sobre un conjunto de obleas (se muestra la introducción dentro del horno de oxidación en una "caja" de cuarzo)** *(derecha, centro);* **serrado de surcos en la superficie de la oblea antes de romperse ésta en distintos dados** *(izquierda, abajo);* **rotura de la oblea en dados, cada uno de los cuales contiene un circuito completo** *(derecha, abajo).* **Las tres primeras fotografías han sido realizadas en la planta de Siltec Cp en Menlo Park, California. Las tres últimas lo han sido en el Centro Zilog. En la construcción de las distintas capas del circuito intervienen otras fases del proceso, no mostradas aquí.**

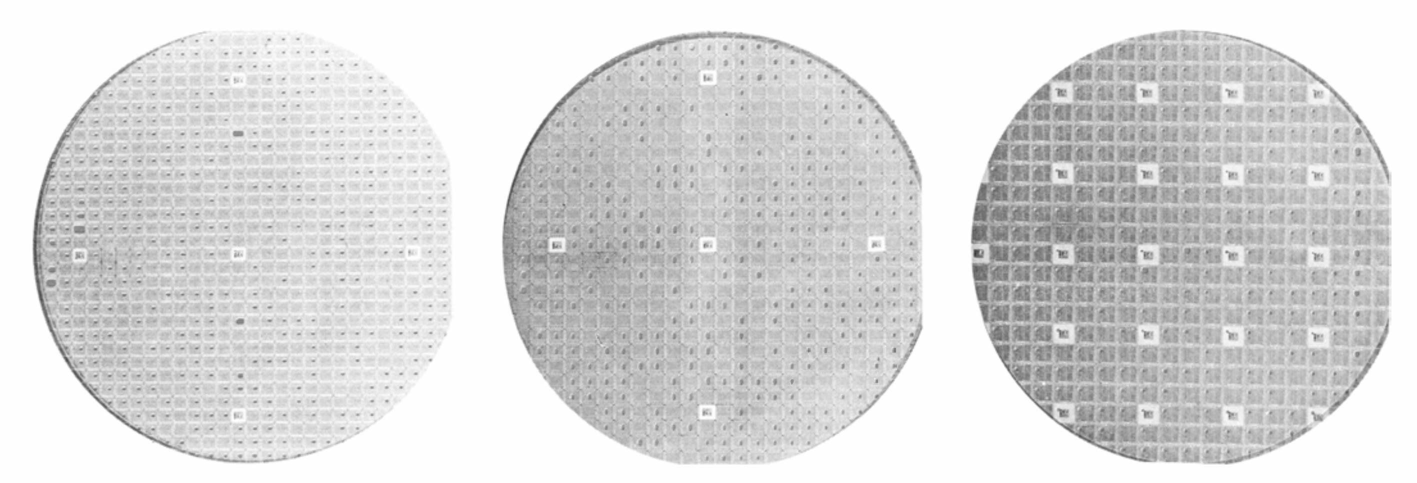




LA COMPROBACION ELECTRICA, para distinguir los circuitos buenos de los defectuosos, se realiza antes de que la oblea de silicio se rompa en dados sueltos. Para este fin, se usa una tarjeta de prueba apropiada, con una serie de contactos parecidos a agujas extremadamente finas. La máquina de comprobación, controlada por ordenador, se ubica por sí sola sobre cada lado, establece contactos con las pistas de circuito, comprueba el circuito y, si el circuito es defectuoso, lo marca con un punto de tinta para indicar que debe desecharse. La máquina salta entonces al siguiente dado y repite la operación. La foto de la izquierda muestra un operador supervisando la comprobación a través de un microscopio. Por su parte, la fotografía de la derecha recoge un primer plano del circuito que se encuentra sometido a comprobación.



UN SOLO DEFECTO puede arruinar todo un microcircuito electrónico. Esta micrografía por barrido electrónico muestra una pequeña porción de una memoria de 16.384 bits. Una partícula de polvo en la máscara empleada para definir la capa de aluminio ha causado la rotura de una conexión eléctrica. Las tiras de aluminio tienen 6 micras (milésimas de milímetro) de ancho. La ampliación es a la escala 1600:1.



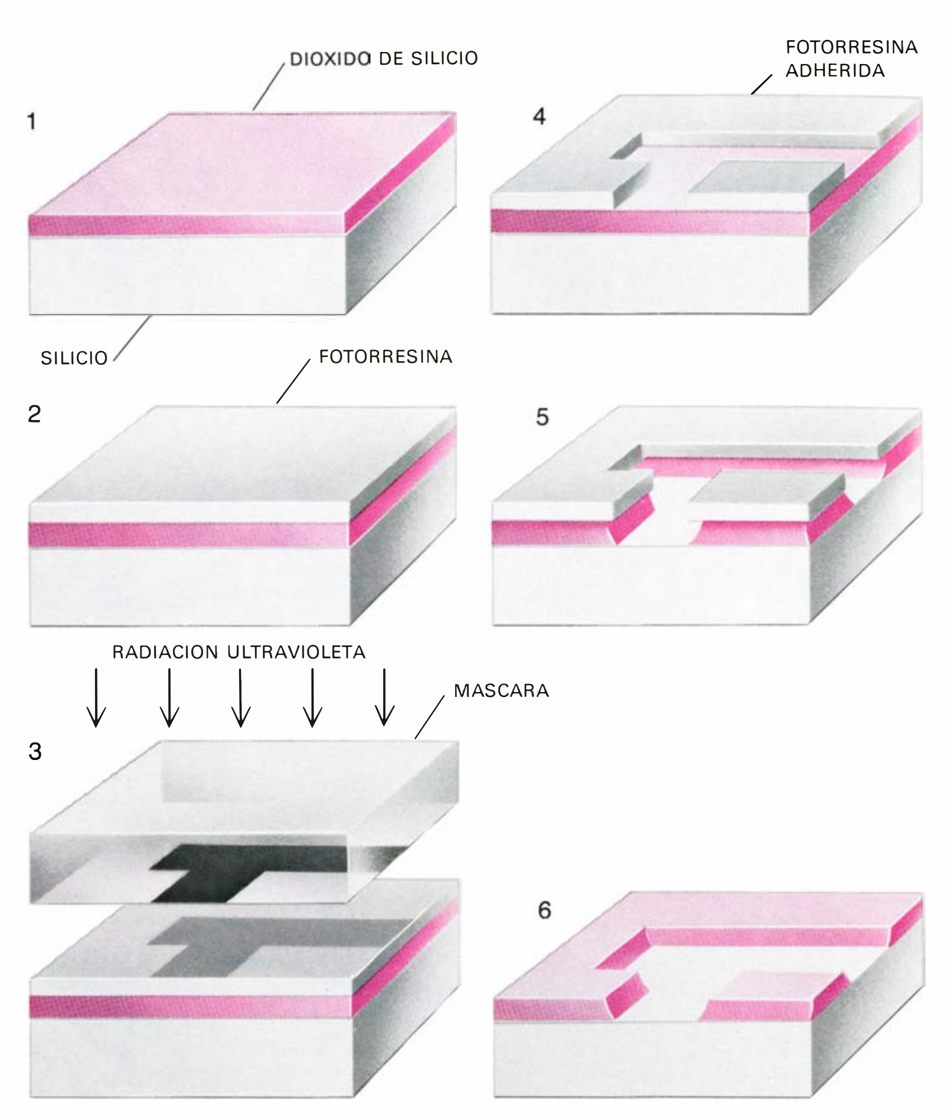
EL EFECTO del tamaño del circuito en el rendimiento se demuestra con tres microcircuitos de dimensiones diferentes. Las obleas que contienen los circuitos miden 100 milímetros de diámetro cada una, y todas han sido fabricadas en el mismo centro de producción de Zilog. Se han sometido a comprobación los circuitos de las obleas y se han marcado los dados defectuosos con un punto de tinta. La oblea de la izquierda contiene 340 circuitos buenos, cada uno de 10,2 milímetros cuadrados de superficie. La oblea del centro contiene 240 circuitos buenos, cada uno de 13,7 milímetros cuadrados de superficie. La oblea de la derecha contiene 80 circuitos buenos, cada uno de 22,9 milímetros cuadrados de superficie. (En la práctica, el rendimiento varía mucho de una oblea a otra incluso para el mismo tipo de circuito; estas tres obleas han sido seleccionadas por ser representativas de las tendencias generales.) El número de circuitos buenos por oblea disminuye al aumentar el tamaño del circuito, porque hay menos espacio libre en la oblea y porque cuanto mayor sea el dado, mayor será el número de defectos.

deben ampliarse mucho; normalmente, los diagramas son hasta 500 veces la dimensión del tamaño final del circuito.

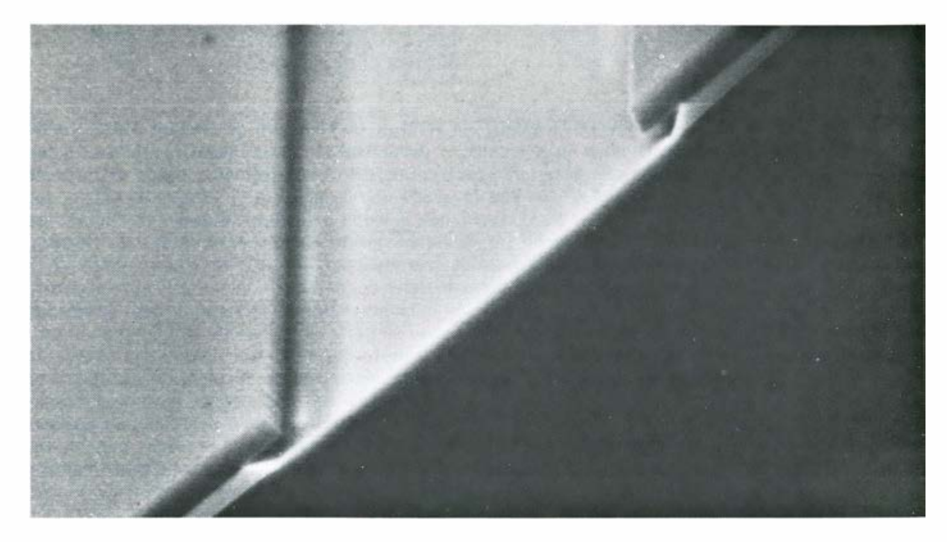
El tiempo requerido para llevar a cabo el diseño y distribución del circuito varía mucho con la naturaleza de éste. Los circuitos más complicados de diseñar son los microprocesadores, cuyo diseño y estratificación puede importar varios años. Otros dispositivos, como las memorias estáticas, con un patrón muy repetitivo pueden ser diseñados y estratificados mucho más rápidamente; en algunos casos bastan unos meses.

Cuando se ha completado el diseño y la estratificación o distribución de un nuevo circuito, la memoria del ordenador contiene una lista de las posiciones exactas de cada elemento del circuito. De esta descripción en la memoria del ordenador se prepara un conjunto de placas, llamadas máscaras. Cada máscara contiene el patrón de una sola capa del circuito. Dado que los circuitos son tan pequeños, se pueden fabricar muchos al mismo tiempo, uno junto a otro, en una sola oblea de silicio. De esta forma, cada máscara, que es una placa de vidrio de unos doce centímetros de lado, tiene un solo patrón, repetido muchas veces en su superficie.

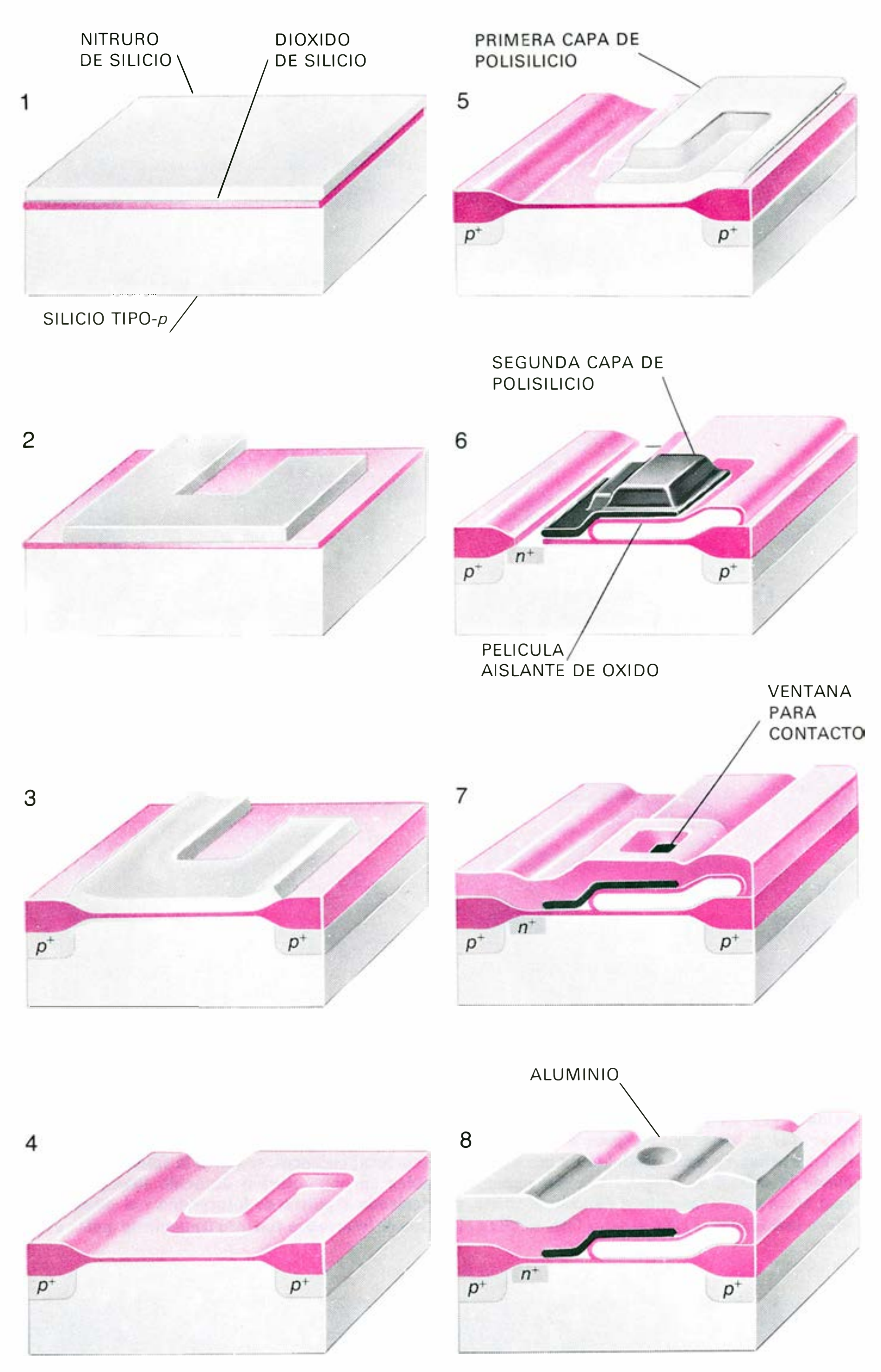
La fabricación de las máscaras es una historia interesante por sí misma. Típicamente el proceso consiste en generar un patrón completo para cada capa del circuito a partir de la memoria del ordenador. Esto se realiza barriendo una placa fotográfica con un haz luminoso controlado por ordenador. Se somete a comprobación este primer patrón, llamado retícula, para detectar sus errores y corregirlo o regenerarlo hasta su perfección. La retícula es 10 veces el tamaño final del circuito. Luego, se proyecta ópticamente sobre la máscara final una imagen de la retícula, que es un décimo de su tamaño original. La imagen se reproduce, de lado a lado, cientos de veces en un proceso iterativo. Se deben superar limitaciones en el sistema mecánico y en el fotográfico; cada elemento debe ser exacto en tamaño y posición con una precisión micrométrica. Se copia, por contacto directo, la placa original, creada con una cámara repetidora, para producir una serie de placas submaestras. Cada submaestra sirve, a su vez, para producir un gran número de réplicas, llamadas placas de trabajo, las cuales se emplearán en el proceso de fabricación. La placa de trabajo puede ser una imagen fijada en una emulsión fotográfica ordinaria, o un patrón mucho más duradero, grabado en una película de cromo sobre un substrato de vidrio.



FOTOLITOGRAFIA es el proceso mediante el cual un patrón microscópico se transfiere de una máscara a una capa de material en un circuito real. En esta ilustración se muestra un patrón que se graba en una capa de dióxido de silicio *(color)* en la superficie de una oblea de silicio. La oblea oxidada *(1)*, en primer lugar, se recubre con una capa de un material fotosensible llamado fotoresina *(2)*, y luego se la expone a la luz ultravioleta a través de la máscara *(3)*. Tras la exposición, la fotorresina se hace insoluble en una solución reveladora; entonces, se retira la fotorresina de allí donde la máscara es opaca *(4)*. Se sumerge la oblea en una solución de ácido fluorhídrico, el cual ataca selectivamente al dióxido de silicio, manteniendo el patrón de fotorresina y el substrato de silicio inalterados *(5)*. En el paso final se quita el patrón de fotorresina con otro tratamiento químico *(6)*.



LA ESTRUCTURA BICAPA, correspondiente al paso 5 de la ilustración superior, se pone de manifiesto a través de una micrografía por barrido electrónico. La muestra enseña la sección recta justo después de que la capa de óxido quedara grabada. El silicio aparece negro. Ampliación 10.000:1.



LA SECUENCIA COMPLETA DE FABRICACION de un elemento de circuito MOS (óxido metálico semiconductor) de canal *n* (que significa: portadores de carga negativa), binivel y con puerta de polisilicio, requiere seis pasos de enmascaramiento. Los primeros pasos del proceso implican la oxidación selectiva del silicio con la ayuda de una película de nitruro de silicio, que actúa como máscara de oxidación. Se hace crecer una delgada película de dióxido de silicio sobre toda la oblea, y se deposita una capa de nitruro de silicio a partir de un vapor químico *(1)*. Se quita selectivamente la capa, en un paso fotolitográfico convencional, de acuerdo con el patrón de la primera máscara *(2)*. Se implanta un contaminante de tipo *p* (por ejemplo, boro) utilizando el nitruro de silicio como máscara, seguido por un paso de oxidación, resultando una gruesa capa de dióxido de silicio en las superficies no enmascaradas *(3)*. Luego se quita el nitruro de silicio en un ataque selectivo que no afecta ni al silicio ni al dióxido *(4)*. Puesto que se consume silicio en el proceso de oxidación, la gruesa capa de óxido está parcialmente sumergida en el substrato de silicio. Se deposita luego la primera capa de silicio policristalino y se recorta en el segundo paso de enmascaramiento *(5)*. Se hace crecer o se deposita una segunda película de óxido, depositando a continuación una segunda capa de silicio policristalino, que es a su vez recortado en el tercer paso de enmascaramiento *(6)*. En esta etapa, un corto ataque con ácido fluorhídrico expone ciertas regiones a una implantación o difusión de contaminante de tipo *n*. A continuación, se deposita una gruesa capa de dióxido de silicio, y se abren "ventanas" para los contactos con la cuarta máscara *(7)*. Finalmente se deposita una capa de aluminio y se recorta en la quinta operación de enmascaramiento *(8)*. La oblea también recibirá una sobrecubierta protectora de dióxido de silicio o nitruro de silicio (no mostrado); el hecho de que se deba proveer de aberturas a esta sobrecubierta en los puntos de soldadura aumenta hasta seis el número de fases de enmascaramiento a recorrer.

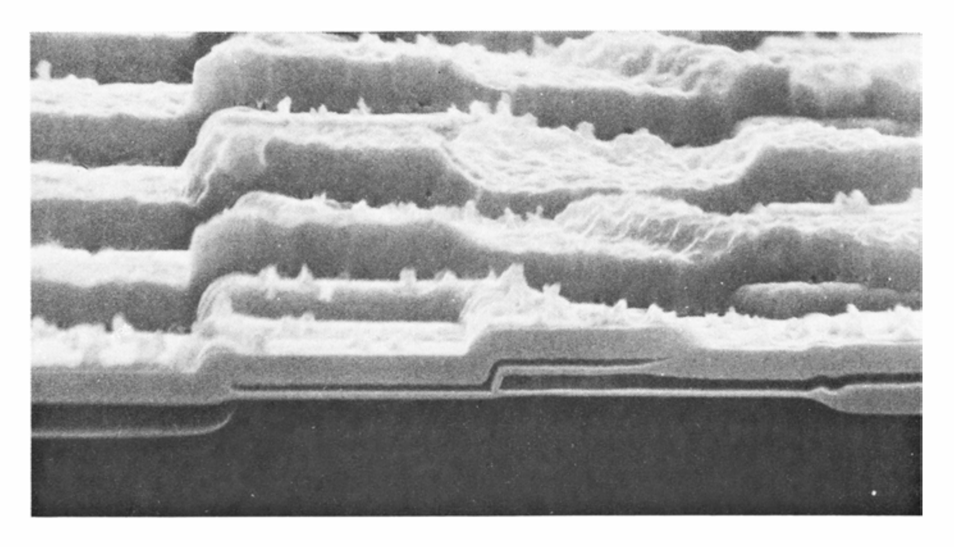
Completado el conjunto de máscaras correctas, culmina la fase de diseño del desarrollo del circuito microelectrónico. Las placas se envían al centro de fabricación de obleas, en donde se emplearán para producir la secuencia deseada de patrones en una estructura física. Este centro de manipulación recibe obleas de silicio, productos químicos y máscaras. Un pequeño centro típico que emplee 100 personas puede procesar varios miles de obleas en una semana. Suponiendo que hay 50 circuitos buenos por oblea, esa planta puede producir cinco millones de circuitos al año.

El interior del centro de fabricación de obleas debe ser extremadamente limpio y ordenado. A causa de la pequeñez de las estructuras fabricadas, no se puede tolerar ni la más pequeña mota de polvo. Una partícula de polvo puede causar un defecto, que tendrá como consecuencia el funcionamiento incorrecto de un circuito. Se llevan trajes especiales para proteger, del polvo transportado por los cuerpos de los operadores, la sala de fabricación. El aire se filtra y recicla continuamente para mantener al mínimo el nivel de polvo. Contando todas las partículas de polvo de un micra o más de diámetro, una planta de fabricación de obleas típica alberga menos de 3 partículas por decímetro cúbico. A efectos de comparación, el nivel de polvo en un moderno hospital es del orden de 300 partículas por decímetro cúbico.

A menudo, el fabricante del circuito compra obleas de silicio preparadas y listas para el primer paso de fabricación. El bajo precio de una oblea preparada (menos de 10 dólares) no justifica las dificultades que entraña su fabricación. Primeramente se reduce silicio puro a partir de su óxido, el componente principal de la arena común. Sigue una serie de procesos químicos para purificarlo hasta que el nivel de pureza alcanza el 99,9999999 por ciento. Una carga de silicio puro se coloca en un crisol y se calienta hasta su punto de fusión: 1420 grados Celsius. Es necesario mantener una atmósfera pura de gas inerte sobre el silicio mientras se funde, para evitar la oxidación y las impurezas indeseadas. Las impurezas deseadas, conocidas como contaminantes, se añaden al silicio para conseguir un tipo de conductividad específica, caracterizada por tener portadores con carga positiva (tipo-*p*) o negativa (tipo-*n*).

Se hace crecer de la fundición un gran monocristal insertando un monocristal perfecto, que hace las veces de "semilla". Girando y tirando lentamente de ella se pueden extraer de la fundición mono-

cristales de siete a diez centímetros de diámetro y varios metros de largo. Se recorta la superficie irregular del cristal que se ha hecho crecer, para obtener un cilindro de diámetro normalizado, de unos setenta y seis o cien milímetros. Se monta el cristal en una mordaza y se corta en obleas con una rápida y fina sierra de diamante. En la última etapa, las obleas son primeramente pulidas por las dos caras y después son rectificadas (finamente pulidas) por una cara. La oblea final tiene un espesor típico de medio milímetro. Los pasos finales se deben llevar a cabo en un ambiente absolutamente limpio. En la superficie acabada no puede haber defectos, desperfectos en el rectificado, arañazos ni siquiera impurezas químicas.



SECCION RECTA del microcircuito electrónico de una memoria de 16.384 bits fabricada mediante el método MOS de canal *n*, binivel y puerta de polisilicio. La estructura corresponde a la etapa final de la ilustración anterior. Esta micrografía fue obtenida en microscopio de barrido.

El papel dominante del silicio como material para microcircuitos electrónicos es atribuible en gran parte a las propiedades de su óxido. El dióxido de silicio es un cristal claro, que funde por encima de los 1400° C. Desempeña un papel muy importante en la fabricación de dispositivos de silicio y en su funcionamiento. Si se calienta una oblea de silicio en una atmósfera de oxígeno o vapor de agua, se forma una película de dióxido de silicio en su superficie. La película es dura y resistente, se adhiere bien y constituye un excelente aislante. El dióxido de silicio es importante en la fabricación de circuitos integrados porque puede actuar como máscara para introducir contaminantes selectivamente. Se puede hacer crecer el espesor de dióxido de silicio adecuado, a temperaturas comprendidas entre 1000 y 1200° C. El espesor exacto puede controlarse precisamente mediante una apropiada selección de la duración y la temperatura de la oxidación. Por ejemplo, una capa de óxido de un espesor de una décima de micra tardará una hora en crecer a una temperatura de 1050° C, en una atmósfera de oxígeno puro. En vapor de agua, a la misma temperatura y en el mismo tiempo, crecerá una capa cinco veces más gruesa.

Un aspecto a tener en cuenta en el proceso de oxidación es su bajo coste. Varios cientos de obleas pueden oxidarse simultáneamente en una sola operación. Las obleas se colocan en ranuras en una "caja de cuarzo", separadas por sólo unos pocos milímetros. El horno de alta temperatura tiene un elemento calefactor cilíndrico, que rodea un largo tubo de cuarzo. A través del tubo pasa un vapor purificado de gas de oxígeno. Las cajas de obleas se colocan en el extremo abierto y son empujadas lentamente hacia la parte caliente del horno. La temperatura en la zona de proceso se controla con precisión mayor que un grado centígrado. El proceso completo suele ser supervisado por un ordenador. Un pequeño ordenador de control de procesos registra la temperatura, dirige la inserción y la extracción de las obleas y controla el ambiente interno del horno.

La fabricación de circuitos integrados requiere un método para formar patrones en la oblea que sea muy preciso. El microcircuito electrónico se construye capa a capa, recibiendo cada capa un patrón a partir de la máscara prescrita en el diseño del circuito. Para conseguirlo, se emplea el proceso de fotograbación conocido como fotolitografía, o simplemente enmascaramiento.

El paso más elemental del proceso de enmascaramiento implica la grabación de un patrón en el óxido. Primeramente se protege una oblea oxidada con una fotorresina, un polímero fotosensible. A este propósito, se vierte en la oblea una gota de fotorresina, diluida en un disolvente, y a continuación se la hace girar rápidamente. Se produce así una delgada película de líquido sobre la superficie, y el disolvente se evapora, dejando la película del polímero. Se aplica un suave tratamiento de calor para secar completamente la película y para aumentar su adhesión a la capa de dióxido de silicio que hay bajo ella.

La propiedad más importante de la fotorresina es que su solubilidad en ciertos disolventes está muy afectada por la exposición a la radiación ultravioleta. Por ejemplo, una fotorresina negativa reacciona y forma polímeros cuando se expone. Por tanto, una exposición a través de una máscara, seguida de un revelado (lavando con el disolvente selectivo), tiene como consecuencia la desaparición de la película allí donde la máscara era opaca. El patrón de fotorresina se calienta después del revelado para endurecerlo.

La oblea, con su patrón de fotorresina, se coloca ahora en una solución de ácido fluorhídrico. El ácido disuelve la capa de óxido allí donde no está protegida, pero no ataca a la fotorresina, ni a la oblea de silicio, propiamente dicha. Una vez que el ácido ha retirado el óxido de silicio de las áreas expuestas, se aclara y se seca la oblea; el patrón de fotorresina se quita con tratamiento químico.

Se recortan otras películas de una forma similar. Por ejemplo, una solución tibia de ácido fosfórico ataca selectivamente al aluminio; por tanto, puede servir para recortar una película de aluminio. A menudo se necesita una capa enmascaradora intermedia, cuando la fotorresina no puede soportar un ataque de una solución grabadora en particular. Por ejemplo, a menudo se graban películas de silicio policristalino con una mezcla particularmente corrosiva que contiene ácido nítrico y ácido fluorhídrico. En este caso, primeramente se hace crecer una película de dióxido de silicio en el silicio policristalino. El dióxido de silicio se recorta de manera normal y se quita la fotorresina. El patrón en el dióxido de silicio puede ahora servir como máscara, para grabar la película de silicio que hay bajo él, porque el óxido sólo es atacado muy lentamente por la mezcla.

Por muchas razones, la fotolitografía encierra la llave de la tecnología microelectrónica. Se la usa repetidamente en el procesado de cualquier dispositivo, una vez al menos por cada capa de la estructura final. Un requisito importante del proceso fotolitográfico es que cada pa-

trón debe colocarse con mucha precisión en relación con las capas que yacen bajo él. Una técnica consiste en colocar la máscara justo encima de la superficie y alinearla visualmente con los patrones de la oblea. La máquina que soporta la oblea y la máscara puede ajustarse en esta operación con una precisión de una o dos micras. Después de que se ha conseguido un alineamiento perfecto, se pone la máscara en contacto con la oblea. Luego, la máscara se ilumina con radiaciones ultravioleta para exponer la fotorresina. A menudo se hace el vacío en el espacio que media entre la oblea y la máscara para conseguir un contacto íntimo; la presión atmosférica comprime la máscara y la oblea, una contra otra. Según se emplee un vacío alto o un vacío moderado, el proceso se llama impresión por contacto "duro" o "blando". En otra variante, impresión por proximidad, la máscara resta sostenida ligeramente por encima de la oblea durante la exposición.

Las variaciones en el proceso de enmascaramiento surgen de la necesidad de imprimir en el patrón marcas muy pequeñas, sin defectos. Si la máscara se colocase muy lejos de la superficie, la difracción de las radiaciones ultravioleta al pasar a través de la máscara provocaría el emborronamiento de las marcas pequeñas. Por ello se prefiere el contacto duro. Por otra parte, la presencia de motas en la oblea o en la máscara rayan la máscara cuando ésta se comprime contra la oblea. Por esta razón, las máscaras sólo se pueden emplear para unas pocas exposiciones, antes de que los defectos acumulados adquieran un nivel no tolerable. Se escoge la técnica de enmascaramiento que mejor se conforme con la tecnología empleada. Dependiendo de las irregularidades de la oblea y de la máscara, y del tipo de máscara empleado, se escoge una técnica que dé una razonable vida a la máscara y una resolución suficiente para imprimir los elementos más pequeños del circuito en el dispositivo.

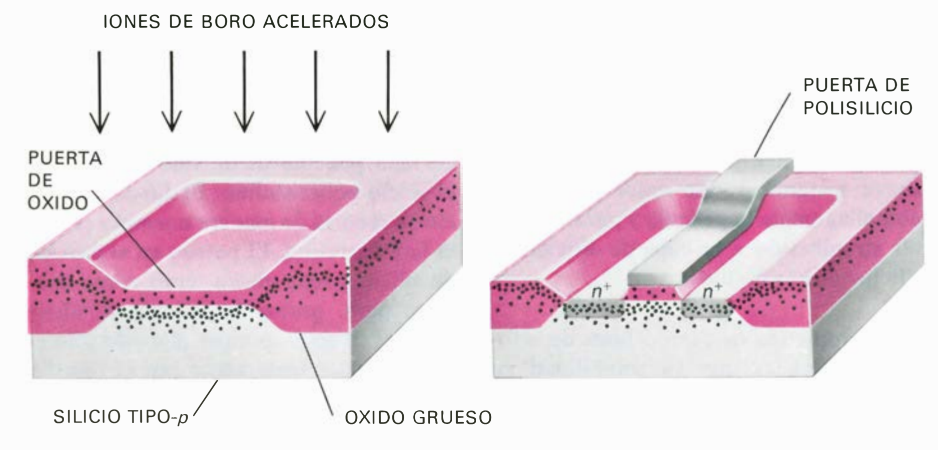
Una reciente tendencia apunta hacia la técnica conocida como alineación por proyección, en la cual la imagen de la máscara se proyecta en la oblea a través de un sistema óptico. En este caso, la vida de la máscara es virtualmente ilimitada. Sólo hace escasos años que se puede disponer de una óptica capaz de cumplir los requisitos de la fotolitografía para la fabricación de circuitos integrados. El hecho de que las obleas aumenten de tamaño cada pocos años es un continuo problema, y el trabajo de diseñar ópticas capaces de formar una imagen precisa sobre la gran área se hace cada vez más difícil. En cualquier caso, recientes alineadores por proyección están superando la grave dificultad de conseguir lentes capaces de resolver detalles del tamaño de una micra en áreas de varios centímetros cuadrados. Se ilumina una zona mucho más pequeña, del orden de un centímetro cuadrado; la exposición se repite haciendo saltar o desplazando la imagen en la oblea.

Los elementos activos del circuito, como transistores de óxido metálico semiconductor (MOS) o transistores bipolares, se forman en parte dentro del substrato de silicio. Para construir estos elementos es necesario introducir impurezas de una forma selectiva, esto es, crear regiones localizadas tipo-*n* y tipo-*p*, por la adición de átomos contaminantes adecuados. Hay dos técnicas para introducir contaminantes selectivamente en el cristal de silicio: difusión e implantación de iones.

Si se calienta el silicio a alta temperatura, a 1000º Celsius, por ejemplo, los átomos de impurezas comienzan a moverse lentamente a través de cristal. Ciertas impurezas clave (boro y fósforo) se mueven mucho más lentamente a través del dióxido de silicio que a través del silicio propiamente dicho. Este hecho importante nos permite emplear finos patrones de dióxido como máscaras para impurezas. Por ejemplo, se puede colocar una caja de obleas en un horno a 1000 grados en una atmósfera que contenga fósforo. El fósforo entra en el silicio por donde no está protegido, difundiéndose lentamente en la masa de la oblea. Una vez se han acumulado suficientes átomos de impurezas, las obleas son retiradas del horno y la difusión de estado sólido cesa inmediatamente. Por supuesto, cada vez que se recalienta la oblea, las impurezas comienzan a difundirse de nuevo; por ello en el diseño de un proceso, para conseguir una profundidad de difusión específica, hay que tener en cuenta todos los tratamientos previstos. Las variables importantes que controlan la profundidad a la cual las impurezas se difunden son el tiempo y la temperatura. Por ejemplo, durante una hora a 1100 grados puede difundirse una capa de fósforo de una micra de profundidad.

Para conseguir un control máximo, la mayoría de las difusiones se hacen en dos pasos. El primer paso, o predepósito, tiene lugar en un horno cuya temperatura se selecciona para conseguir el mejor control posible de la cantidad de impurezas introducidas. La temperatura determina la solubilidad del contaminante en el silicio, exactamente igual que la temperatura del agua caliente determina la solubilidad de una impureza como la sal. Después de un tratamiento de depósito relativamente corto, la oblea se coloca en un segundo horno, normalmente a mayor temperatura. Este segundo tratamiento con calor, el paso de "difusión conducida", se controla para conseguir la profundidad de difusión deseada.

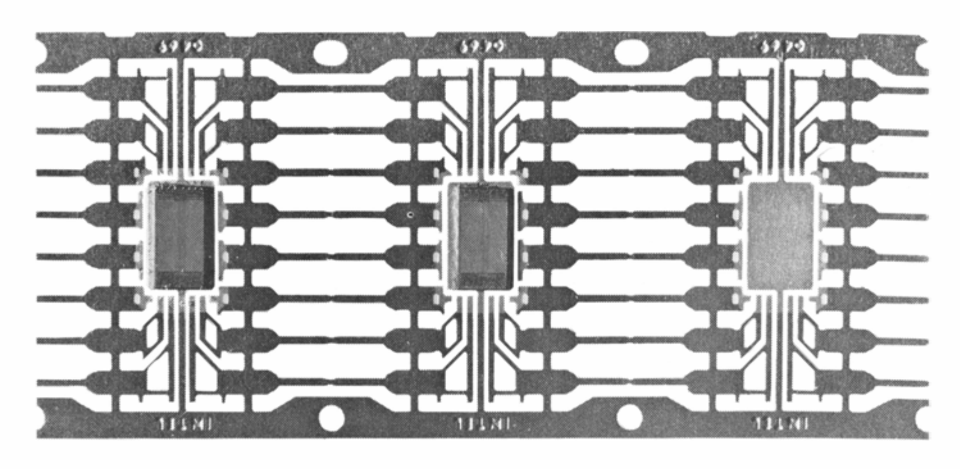
En la formación de uniones *pn* por difusión de estado sólido, las impurezas se difunden lateralmente bajo la máscara de óxido, la misma distancia, más o menos, que la profundidad de la unión. El flanco de la unión *pn* queda entonces protegido por una capa de dióxido de silicio. Esta es una característica importante de la técnica, porque el dióxido de silicio es casi un aislante ideal, y muchos dispositivos electrónicos no toleran ninguna filtración en el flanco de la unión.

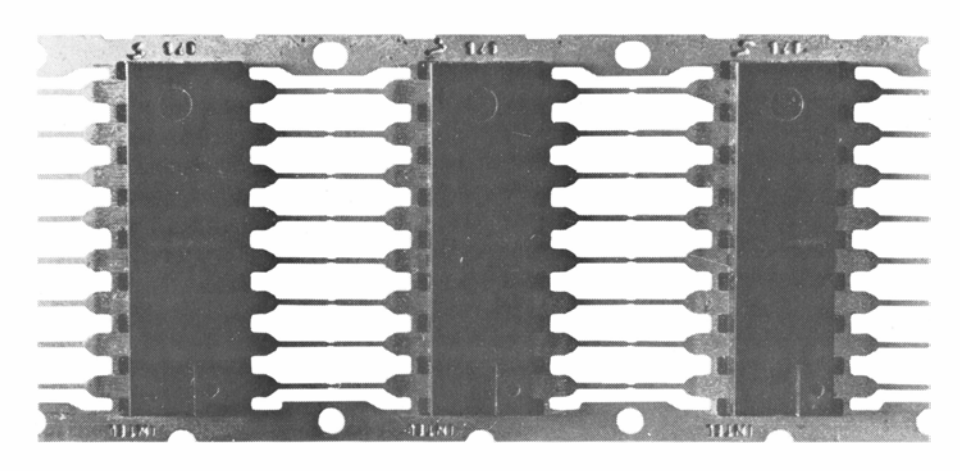


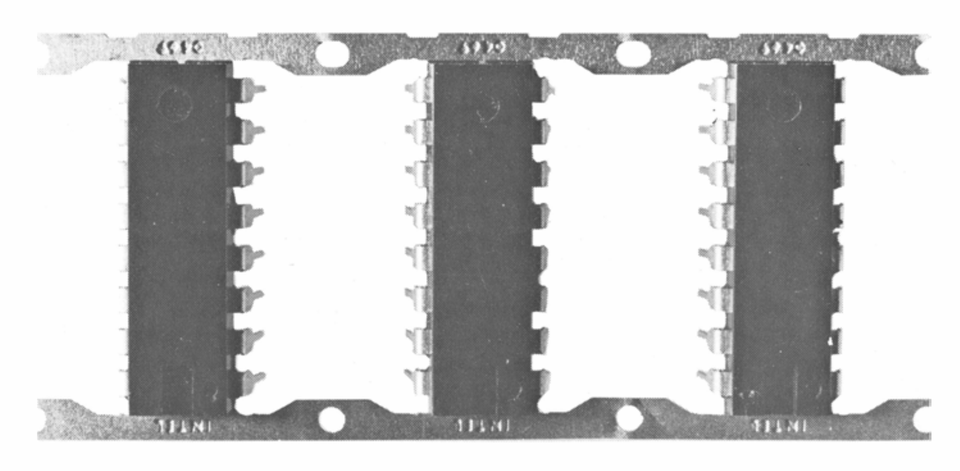
IMPLANTACION DE IONES, para controlar con precisión la cantidad de contaminante (en este caso iones boro) bajo la puerta de óxido de un transistor MOS. Con una tensión de aceleración apropiada, se puede conseguir que los iones atraviesen la puerta de óxido pero no el óxido más grueso *(izquierda)*. Una vez implantados los iones boro, el silicio policristalino se deposita y recorta para formar las regiones de la puerta del transistor. Se quita luego una fina capa del óxido y se forman las regiones del surtidor y drenador del transistor por difusión de impurezas de tipo *n (derecha)*.

Se ha desarrollado otro proceso de contaminación selectiva, la implantación de iones, mediante la introducción de impurezas a temperatura ambiente. Los átomos contaminantes son ionizados (privados de uno o más de sus electrones) y acelerados hasta conseguir gran energía, haciéndolos pasar a través de una diferencia de potencial de decenas de miles de voltios. Al final de su recorrido chocan con la oblea de silicio y se incrustan a profundidad variable, dependiendo de su masa y su energía. La oblea se puede enmascarar frente a los iones mediante una capa de dióxido recortada, que actúa de máscara como en la difusión convencional, o por un patrón de fotorresina. Por ejemplo, los iones fósforo acelerados por un potencial de 100.000 voltios penetrarán en la fotorresina a una profundidad de menos de media micra. Allí donde chocan con silicio descubierto penetran hasta una profundidad de una décima de micra, por término medio. Por tanto, incluso una capa de fotorresina de una micra puede servir de máscara para la implantación selectiva del fósforo.

Dado que los iones acelerados perforan su camino en el cristal de silicio, causan daños considerables en la estructura del cristal. De todas formas, es posible reparar la mayoría de los daños templando el cristal a una temperatura moderada. A esta temperatura, se produce una pequeña difusión; teniéndolo en cuenta, se pueden escoger las condiciones de implantación de iones que producirán la distribución deseada. Por ejemplo, se puede conseguir una concentración de contaminantes alta y muy poco profunda con una adecuada implantación de iones. La característica más significativa de la técnica es la posibilidad de controlar con precisión la concentración de contaminantes. Los iones que bombardean el cristal transportan una carga cada uno, y se puede determinar exactamente el número de impurezas midiendo la carga total que se acumula. Por ello, la implantación de iones se usa siempre que el nivel de contaminantes se tiene que controlar con precisión. A menudo, la implantación de iones sustituye simplemente el paso de predepósito del proceso de difusión. La implantación de iones también se emplea para introducir impurezas que son difíciles de predepositar a partir de un vapor a alta temperatura. Por ejemplo, las investigaciones sobre el uso de arsénico para formar una fina capa de contaminante tipo-*n* en dispositivos MOS coinciden con la posibilidad de emplear equipos de implantación de iones.









ENCAPSULADO o etapa final de la confección de un microcircuito electrónico, representado en esta secuencia de fotografías. En primer lugar, los dados que contienen los circuitos acabados se sujetan a un metal estampándolos. Se conectan finos hilos metálicos desde los puntos de soldadura de los dados a los electrodos de la cápsula; se moldea una cubierta de plástico alrededor de cada dado y se separan las unidades. Los componentes retratados fueron aportados por Intel Co.

Una característica única de la implantación de iones es la posibilidad de introducir impurezas a través de un fino óxido. Esta técnica es particularmente ventajosa en el ajuste de la tensión umbral de los transistores MOS. Contaminantes tipo-*n* o tipo-*p* pueden implantarse a través de la puerta de óxido, lo que provoca una disminución o un aumento de la tensión umbral del dispositivo. Por tanto, mediante la técnica de implantación de iones se pueden fabricar varios tipos distintos de transistores MOS en la misma oblea.

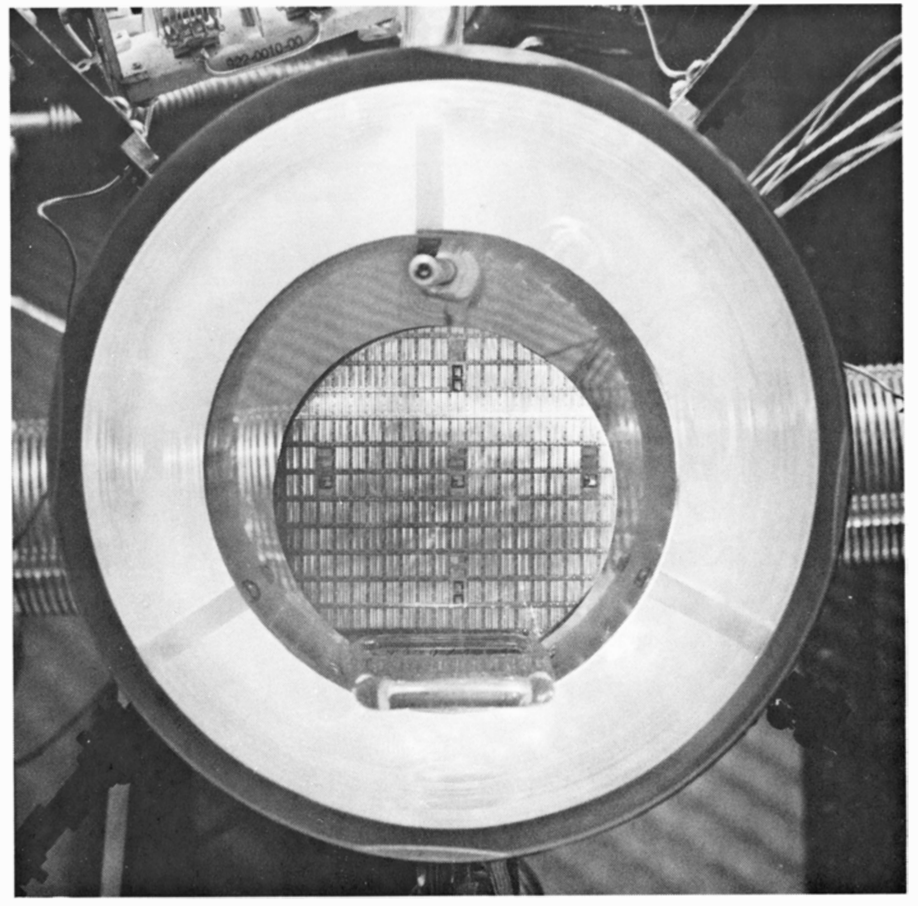
Las capas superiores de los circuitos integrados se forman depositando y recortando finas películas. Los dos procesos más importantes para depositar finas películas son deposición de vapor químico y evaporación. La película policristalina de silicio en la puerta de silicio de la importante tecnología MOS suele colocarse mediante deposición de vapor químico. El gas silano (SiH_4) se descompone cuando se calienta, separándose el silicio y el hidrógeno. De acuerdo con esto, cuando las obleas se calientan en una atmósfera diluida de silano, se forma lentamente en la superficie una película uniforme de silicio policristalino. En pasos posteriores la película se contamina, oxida y recorta.

También se pueden depositar películas aislantes, como dióxido de silicio o nitruro de silicio, mediante depósito de vapor químico. Si una fuente de oxígeno, tal como dióxido de carbono, está presente durante la descomposición del silano, se forma dióxido de silicio. Similarmente, se hace crecer nitruro de silicio descomponiendo silano en presencia de un compuesto de nitrógeno, tal como el amonio.

El método de evaporación es tal vez el más sencillo para depositar una película delgada; se emplea normalmente para colocar la capa de conductor metálico en la mayoría de los circuitos integrados. La carga metálica a evaporar, normalmente aluminio, se coloca en un crisol y las obleas a recubrir se sitúan encima del crisol en una mordaza móvil llamada planetario. Durante la evaporación, las obleas se hacen girar para asegurar la máxima uniformidad de la capa. El movimiento del planetario también hace oscilar las obleas con respecto a la fuente, con objeto de obtener una película continua de aluminio sobre los escalones y baches existentes en la superficie, creados por los pasos fotolitográficos precedentes. Después, se hace bajar una campana de vidrio sobre el dispositivo planetario y se establece un alto vacío: luego, se calienta la carga de aluminio por bombardeo directo con electrones de alta energía. En la oblea, se deposita una película de aluminio puro.

En la fabricación de un circuito integrado a gran escala típico hay más pasos de depósito de película fina que de difusión. Por tanto, la tecnología de película fina probablemente es más crítica, en lo referente al rendimiento total y las prestaciones de los circuitos, que las fases de oxidación y difusión. En un desarrollo reciente se emplea película fina para seleccionar las áreas de la oblea que deben ser oxidadas. El compuesto nitruro de silicio tiene la propiedad de oxidarse mucho más lentamente que el silicio. Una capa de nitruro de silicio puede depositarse al vapor, recortarse y utilizarse como máscara para la oxidación. La superficie resultante es mucho más lisa que la superficie obtenida haciendo crecer óxido por todas partes y retirándolo selectivamente. Los dispositivos MOS de canal *n* cuentan con la ventaja adicional de que se puede añadir una fase de implantación de iones antes de la etapa de oxidación; en la implantación se emplea boro, y el patrón de nitruro actúa de máscara. Este proceso provoca una región de tipo *p* fuertemente contaminada, que se instala bajo el óxido, la cual entorpece la formación de canales desde los elementos adyacentes del dispositivo.

Este paso de difusión del "canal frenador" es necesario en la tecnología MOS de canal *n* de altas prestaciones. Sin la oxidación selectiva, se tendría que añadir una fase de enmascaramiento especial. La separación entre los elementos tendría que ser forzosamente mayor; por tanto, la oxidación selectiva tiende a aumentar la densidad del circuito. Los circuitos integrados bipolares también se benefician en gran manera del uso de la oxidación selectiva. Sustituyendo el aislamiento de difusión convencional por el aislamiento con óxido, el espacio ocupado por el transistor bipolar se reduce a menos de la cuarta parte.



GRABADO SECO POR PLASMA, avanzada técnica de proceso actualmente en desarrollo en el laboratorio del autor en la Universidad de California en Berkeley. La vista está tomada a través de una ventana de cuarzo en un extremo de la cámara de reacción. Moléculas de tetrafluorhidrato de carbono gaseoso (freón), a una presión unas mil veces menor que la atmosférica, fluyen a través de la cámara de cuarzo, donde quedan ionizadas por una descarga de radiofrecuencia. Un escudo de aluminio perforado protege las obleas recubiertas de polisilicio de las descargas directas. Por las superficies de las obleas se difunden átomos libres de flúor, creados en la descarga por rotura de las moléculas de freón, y reaccionan con el silicio policristalino allí donde no está protegido por fotorresina. El proceso produce patrones más precisos que el método de grabado húmedo.

La fase de fabricación de obleas finaliza con una comprobación eléctrica. Cada dado de la oblea se somete a análisis para determinar si funciona o no correctamente. Los dados defectuosos son marcados con un punto de tinta para indicar que deben desecharse. Una máquina de comprobación controlada por ordenador examina rápidamente cada circuito, salta al siguiente y hace el entintado sin intervención humana. También puede elaborar estadísticas sobre el número de circuitos buenos por oblea, la posición y la incidencia relativa de los distintos tipos de fallos. Esta información sirve de ayuda eficaz para encontrar nuevos caminos con el fin de aumentar la producción de circuitos buenos.

El circuito completo debe pasar una última operación: el encapsulado. Debe introducirse en una suerte de caja protectora y ha de tener conexiones con el mundo exterior. Hay muchos tipos de cápsulas, siendo común a todas ellas poseer un tamaño mucho mayor y ser más fuertes que los dados de silicio propiamente dichos. En primer lugar, se secciona la oblea para separar las pastillas individuales, lo que suele hacerse marcando cortes entre las pastillas y rompiendo la oblea a lo largo de las líneas de los surcos. Los circuitos buenos se sueldan a las cápsulas y se conectan mediante finos hilos a los electrodos que salen fuera de la cápsula. La cápsula se sella luego y el dispositivo ya está listo para la comprobación final. Se somete al circuito encapsulado a una serie exhaustiva de comprobaciones eléctricas, para asegurarse de que funciona perfectamente y de que continuará haciéndolo por muchos años.

Una vez se han obtenido las distintas pastillas a partir de la oblea, el coste por cada fase de manipulación aumenta enormemente. El coste ya no se reparte entre muchos circuitos. De acuerdo con esto, se debe introducir manipulación automática durante el encapsulado y la comprobación, siempre que sea posible. La técnica tradicional de ahorro ha consistido en emplear mano de obra barata para la operación de encapsulado. Comoquiera que el coste de la mano de obra aumenta y se puede disponer de mejoras en la tecnología de encapsulado, se está substituyendo gradualmente la mano de obra por el montaje altamente automatizado.

Se encuentran en vías de desarrollo gran número de técnicas de proceso avanzadas. Por ejemplo, el sencillo proceso de grabado húmedo, en el que se eliminan selectivamente partes de las películas de aluminio o silicio policristalino, está siendo sustituido por el proceso de grabado en seco. El silicio policristalino puede ser "grabado por plasma" en un gas de moléculas de tetrafluoruro de carbono (freón) eléctricamente excitado. Una descarga eléctrica de alta frecuencia a baja presión rompe la molécula de freón en varios iones y radicales libres (átomos de flúor, por ejemplo). Los radicales libres atacan a la película pero no reaccionan con la máscara de fotorresina.

Además de ser un método más controlable para eliminar selectivamente el silicio, el grabado por plasma promete ser mucho menos dañino para el ambiente. En vez de producir grandes cantidades de ácidos corrosivos, los productos de la reacción son pequeñas cantidades de flúor y fluoruro de silicio, que son fácilmente captados a la salida del sistema.

La tecnología de la fotolitografía, que se mantuvo estable durante unos diez años, está también sufriendo algunos cambios. Primero, el número de defectos disminuye a medida que la litografía por proyección va sustituyendo a la litografía por contacto. Segundo, la disponibilidad de máscaras de alta calidad está reduciendo constantemente el tamaño y el coste de los dispositivos microelectrónicos. Tercero, se están desarrollando nuevos métodos de litografía para conseguir rebajar a un décimo el tamaño de cada elemento del circuito y reducir a una centésima el área del circuito.

Los detalles más pequeños que se pueden formar con el proceso fotolitográfico convencional están limitados en última instancia por la longitud de onda de la luz. La tecnología actual puede reproducir rutinariamente elementos de sólo algunas micras de diámetro y parece posible reducir los detalles más pequeños hasta aproximadamente una micra. De todas formas, haces de electrones y rayos X tienen longitudes de onda que se miden en nanómetros, y más pequeñas; son, pues, capaces de producir detalles extremadamente finos.

La litografía por rayos X es simplemente una forma de fotolitografía por contacto, en la cual los rayos X débiles sustituyen las radiaciones ultravioletas. La técnica de rayos X ofrece obviamente mayor resolución; se han producido estructuras sencillas de menos de una décima de micra de diámetro. El proceso es potencialmente más rápido y barato, ya que se expone la oblea entera. De todas formas todavía quedan por resolver muchos problemas. Las máscaras para rayos X, que consisten en un patrón metálico sobre una fina membrana (mylar, por ejemplo), son frágiles y difíciles de hacer. También es difícil alinear la máscara con respecto al patrón de la oblea. Debido a la atenuación de los rayos X en el aire, la oblea debe ser expuesta en el vacío o en una atmósfera de helio. Las fuentes actuales de rayos X son comparativamente débiles y se necesitan largas exposiciones. Como resultado de ello, todavía no se han confeccionado circuitos integrados comerciales con la ayuda de la litografía de rayos X.

La litografía con haz de electrones es una técnica más antigua y madura, que tiene su fundamento en el microscopio electrónico. Un sistema para litografías por haz de electrones se parece mucho a un microscopio electrónico de barrido. Un fino haz de electrones barre la oblea para exponer una resina sensible a los electrones en las áreas deseadas. A pesar de que se han recogido impresionantes resultados, la aplicación de la litografía de haz de electrones está limitada por su elevado coste actual. Las máquinas son caras (cerca de un millón de dólares por máquina) y puesto que el haz de electrones debe barrer la oblea en vez de exponerla toda al mismo tiempo, la duración del proceso necesario para poner un patrón en la oblea es bastante largo. De todas formas, la velocidad de los progresos en este campo es rápida y están apareciendo sistemas más prácticos.

A pesar de que actualmente la litografía por haz de electrones resulta demasiado cara como para ser usada en el proceso de fabricación de obleas, ya es una técnica de producción rutinaria en la realización de máscaras fotolitográficas. Con la ayuda del método del haz de electrones, se pueden eliminar dos pasos de reducción fotográfica y escribir el patrón directamente en la máscara a partir de la información almacenada en la memoria del ordenador. De esta forma, se pueden crear máscaras tan sólo unas horas después de haber acabado el diseño. Las ventajas de la alta resolución, simplicidad de manipulación y brevedad del tiempo de producción pueden conducir muy bien a una completa adopción, por parte de la industria, de la realización de máscaras con haz de electrones. Gradualmente, según disminuyan los costes de la litografía de haz de electrones, esta técnica se introducirá en la fabricación de obleas, y nacerá una nueva generación de microcircuitos electrónicos todavía más complejos.

Memorias microelectrónicas

Las memorias actuales, basadas en transistores, suelen almacenar unos 16.000 bits (dígitos binarios) en una pastilla. Los componentes de burbuja magnética y de carga acoplada ofrecen una densidad de almacenamiento de información todavía mayor

David A. Hodges

Los sistemas modernos de procesamiento de la información y de control precisan de un rápido almacenamiento y recuperación de la información digital. La cantidad de información que ha de almacenarse varía desde menos de 100 bits (dígitos binarios) para una calculadora simple de bolsillo hasta un número de bits comprendido entre mil millones y un billón para un ordenador de gran capacidad. La tecnología del almacenamiento digital es quizás el sector de la microelectrónica que cambia más rápidamente. Durante la última década, la velocidad de operación y la fiabilidad han aumentado por lo menos en un orden de magnitud, al tiempo que el tamaño físico, el consumo de potencia y el coste por bit almacenado se han visto reducidos en factores que van de 100 a 1000. En la próxima década se prevén mejoras de magnitud comparable antes de que limitaciones físicas fundamentales fuercen a un ritmo más lento.

En el contexto de la electrónica, el término "memoria" (o, según el uso británico, "almacén") suele referirse a un elemento destinado a almacenar información digital. Las operaciones de almacenaje ("escribir") y de recuperación ("leer") se encuentran por completo bajo control electrónico. El almacenaje analógico de información auditiva o visual se efectúa habitualmente en cinta magnética y se le denomina registro. En el registro analógico, la intensidad de la "banda" en el medio de almacenaje reproduce directamente la intensidad variable de la señal de entrada. Si bien existe cierto solapamiento entre el registro analógico y el digital, me concentraré en este artículo en la memoria digital.

Las memorias digitales más utilizadas son las memorias de lectura-escritura, cuya expresión significa que realizan las operaciones de leer y de escribir a una velocidad idéntica o similar. Es obvio que la capacidad de almacenamiento, el coste por bit y la fiabilidad son características importantes en las memorias. Otras características importantes son la velocidad de operación (definida en términos de tiempo de acceso), la duración del ciclo y la velocidad de transferencia de datos. El tiempo de acceso es simplemente el tiempo necesario para leer o para escribir en cualquier ubicación del almacenamiento. Algunas memorias tales como las memorias de acceso aleatorio (o RAM, del inglés "random-access memory") tienen el mismo tiempo de acceso para cualquier ubicación del almacenamiento. Las memorias de acceso en serie y en bloque poseen tiempos de acceso que dependen de la posición del almacenamiento seleccionada.

Las memorias de acceso aleatorio pueden realizar la totalidad de las operaciobes de leer y de escribir en un período mínimo especificado, conocido como duración del ciclo. Las memorias de acceso en serie y en bloque tienen un tiempo de acceso variable y relativamente grande, transcurrido el cual la velocidad de transferencia de datos es constante. La velocidad de transferencia de datos es aquella velocidad a la que se transfiere la información hacia o desde ubicaciones de almacenamiento secuenciales.

Una clasificación conveniente de las memorias digitales es la que distingue los dispositivos de superficie móvil de los enteramente electrónicos, con subdivisiones en cada categoría. Si bien los dispositivos de superficie móvil no pueden considerarse realmente como microelectrónicos, se están aplicando ventajosamente técnicas microelectrónicas para mejorar su rendimiento. En muchas aplicaciones de las memorias, uno u otro de los dos sistemas, el de superficie móvil o el enteramente electrónico, es claramente superior, pero en ciertas aplicaciones actuales y en otras inminentes ambos sistemas son muy competitivos. Por este motivo describiré brevemente las características de los dispositivos de superficie móvil.

Generalmente, los dispositivos de memoria de superficie móvil realizan el almacenamiento de información en áreas localizadas de una delgada película magnética que recubre una superficie de soporte no magnética. El material magnético debe ser capaz de mantener un flujo magnético remanente en ausencia de campos externos aplicados. La superficie de soporte puede ser flexible, en forma de cinta o disco de plástico, o bien puede ser rígida, en forma de un disco o tambor de aluminio o de algún otro material. La película magnética y la cabeza de lectura-escritura (un pequeño electroimán) se mueven una en relación con la otra a fin de colocar una ubicación del almacenamiento en la posición correcta para efectuar la lectura o la escritura de información. La información se almacena bajo la forma de puntos minúsculos magnetizados en la película magnética. La información almacenada es captada por la cabeza de lectura-escritura en forma de una corriente débil, inducida en ella a medida que la película magnética se mueve bajo la cabeza.

Las memorias digitales de cinta magnética más sencillas proceden de la adaptación de registradores de cinta con "cassettes" de audio. Una "cassette", o un cartucho de cinta, puede almacenar de 10^6 a 10^7 bits y puede proporcionar un tiempo de acceso de 10 a 100 segundos a cualquier ubicación seleccionada del almacenamiento. La velocidad de la cinta está limitada por el desgaste y por el calor generados en el movimiento, ya que la cinta roza con la cabeza (o cabezas) de lectura-escritura.

Existen sistemas digitales de memoria de cinta más elaborados, así como memo-

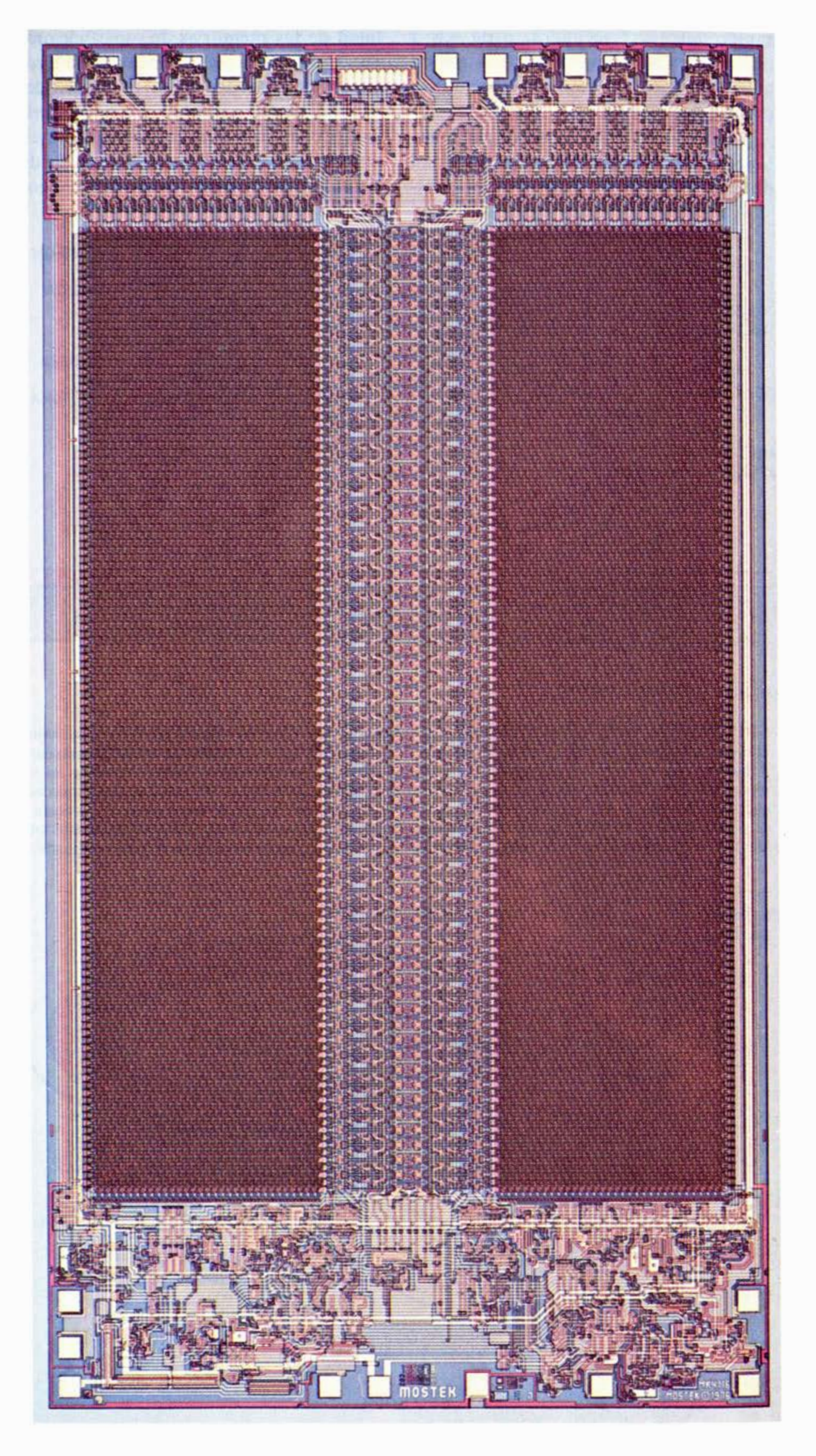
rias de disco y de tambor, que emplean un flujo laminar de aire para mantener la cabeza de lectura-escritura unas pocas micras por encima de la superficie magnética. Esta estratagema hace posible alcanzar velocidades de superficie mucho más altas y mejora la fiabilidad del almacenamiento de datos.

Las memorias de disco magnético de gran capacidad almacenan de 10^9 a 10^{10} bits en una o más superficies de disco, que están montadas en un eje único que gira accionado por un solo motor. El acceso a las posiciones de almacenamiento se consigue moviendo radialmente una o más cabezas de lectura-escritura a través del disco o discos giratorios. En los sistemas típicos, el tiempo promedio de acceso es de 20 milisegundos.

Las memorias de superficie móvil son aparatos de acceso en bloque. Después de un tiempo de acceso determinado por los regímenes del movimiento mecánico y por la posición que se está buscando, los datos se transfieren a una velocidad determinada por la velocidad superficial y por la densidad del almacenamiento de datos en la superficie. Los regímenes de transferencia típicos varían de 10^6 a 10^7 bits por segundo.

La información almacenada en un medio magnético suele conservarse cuando se desconecta la alimentación exterior; tal tipo de almacenamiento se considera como no volátil. En los sistemas de cinta y de disco suele poderse retirar de ellos el medio registrador, con lo cual se hace posible la transferencia física de la información de un lugar a otro y se permite el acceso a cantidades de información virtualmente ilimitadas, por simple intercambio de los discos o carretes de cinta. Estas útiles propiedades no suelen encontrarse en las memorias electrónicas.

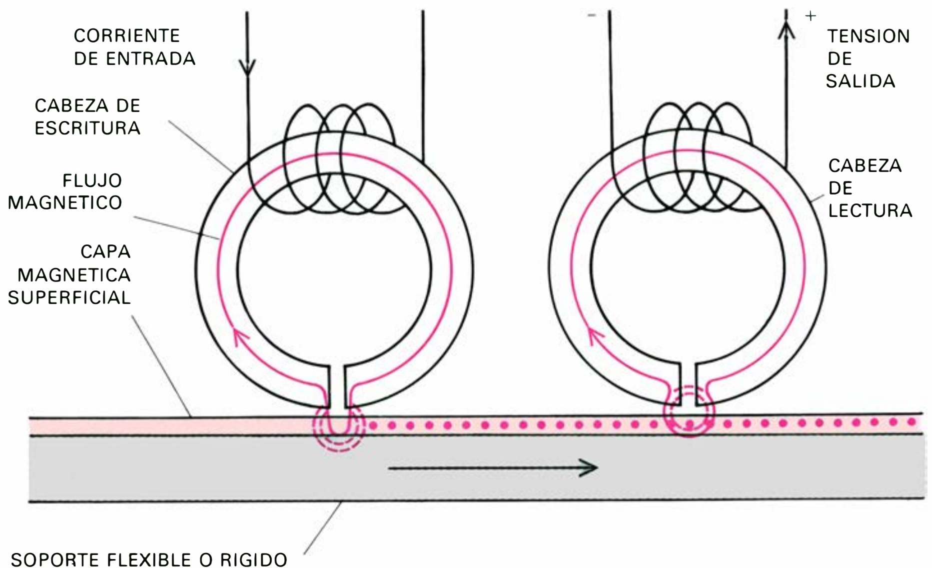
En los sistemas de computadores, los dispositivos de memoria de superficie móvil se utilizan ampliamente como memorias de "archivo", en las que se almacena la información que se cambia mucho menos de lo que se lee. El coste por bit de la información almacenada en las memorias de superficie móvil es más bajo, de uno a cuatro órdenes de magnitud, que el de la información almacenada en las memorias electrónicas (los costes aproximados son de 10^{-5} centavos de dólar por bit para el sistema más barato de cinta magnética contra 0,1 centavos por bit para la memoria microelectrónica menos cara). A menudo, la característica de no volatilidad de las memorias de superficie móvil es esencial. Por otro lado, el tiempo de acceso de las memorias electrónicas puede ser más



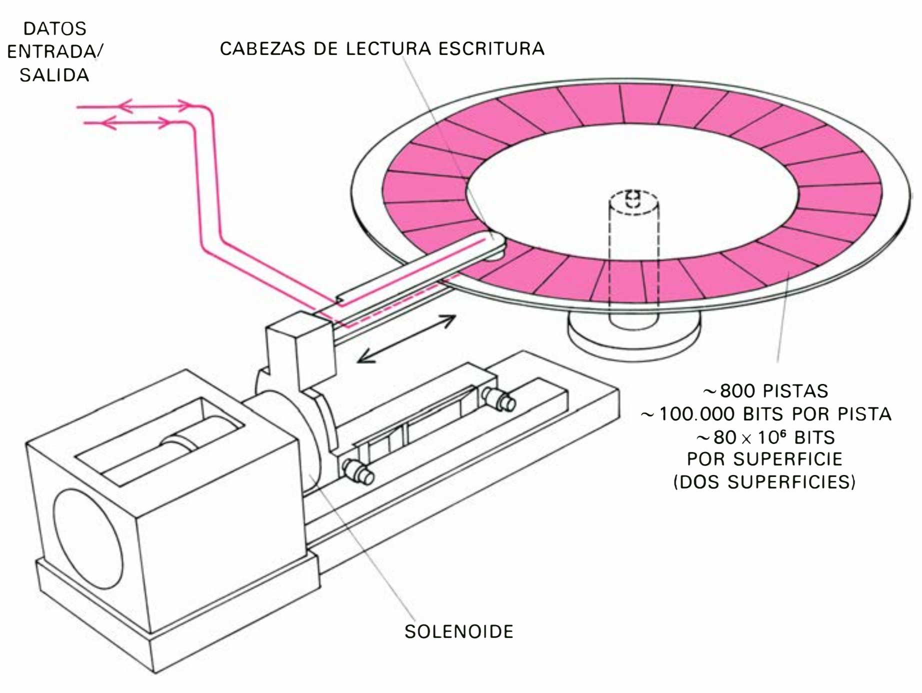
CIRCUITO DE MEMORIA DE ACCESO ALEATORIO (RAM), que proporciona almacenamiento para 16.384 bits (dígitos binarios). Cada bit se mantiene en una célula de almacenamiento de un solo transistor *(véase la ilustración de la página 71)*. **El tiempo requerido para escribir un bit en cualquier posición arbitraria o para leerlo es alrededor de 200 nanosegundos. La pastilla RAM mostrada aquí, que es la MK 4116 fabricada por la compañía Mostek, mide 2,8 por 5,1 milímetros. El componente terminado se vende por 30 dólares en pedidos de 1000 o más unidades.**

rápido que 10^{-8} segundos, o de seis a ocho órdenes de magnitud más rápido que el tiempo de acceso de las memorias de superficie móvil.

Se están realizando continuamente mejoras en el coste por bit, en la velocidad de transferencia de datos y en la fiabilidad de las memorias de superficie móvil. Las técnicas de fabricación microelectrónicas favorecen la reducción del tamaño de las cabezas de lectura-escritura y conducen a la obtención de mayores densidades de bits en la superficie de registro. Los circuitos microelectrónicos sirven también para perfeccionar los códigos de corrección de error en los sistemas de memoria de superficie móvil, reduciendo sustancialmente el ritmo con que se presentan los errores en los procesos totales de almacenamiento y de recuperación de datos. Es claro que, en un futuro previsible y para ciertas tareas esenciales de almacenamiento de información, no es probable que las memorias electrónicas suplanten a los sistemas de superficie móvil.



SUPERFICIE MAGNETICA MOVIL que almacena información binaria al coste más bajo: 100.000 o más bits por un centavo de dólar. La información queda retenida de modo permanente a menos que se borre a propósito. Los dígitos binarios se almacenan secuencialmente en un medio magnético móvil, mediante la aplicación de impulsos eléctricos a un electroimán minúsculo que sirve de cabeza de "escritura". Cuando la película pasa por debajo de una cabeza de "lectura" de diseño similar, puntos magnetizados en la película generan impulsos eléctricos pequeñísimos que luego se amplifican. Una sola cabeza puede servir para leer y para escribir en tiempo compartido.

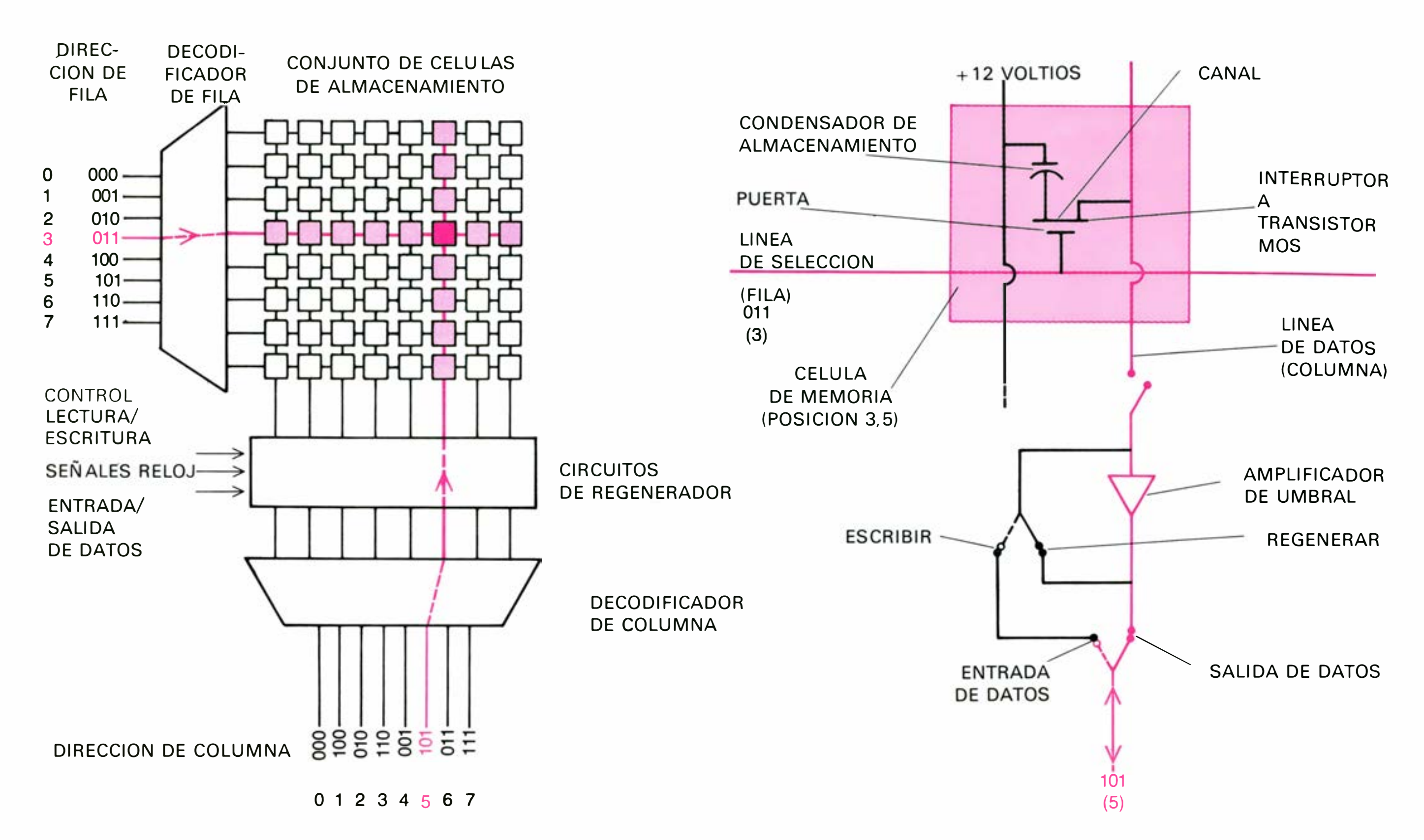


LOS SISTEMAS DE MEMORIA DE DISCO MAGNETICO se emplean ampliamente para ficheros de gran capacidad y consulta frecuente, o como almacenes de información. El sistema simple mostrado aquí sólo tiene un disco, pero son comunes los sistemas con discos múltiples. La información se almacena en ambos lados del disco por cabezas separadas de lectura-escritura. Un sistema moderno típico tiene alrededor de unas 800 pistas por lado con capacidad de 100.000 bits por pista. La capacidad total de almacenamiento de ambos lados de un disco es, por tanto, de 1600 × 100.000, o sea, de 160 millones de bits. Asimismo la información secuencial puede transferirse dentro y fuera del sistema a velocidades de cinco a diez millones de bits por segundo.

Los sistemas más recientes de memoria electrónica han sido posibles gracias a la tecnología moderna de los semiconductores, que permite situar miles de elementos electrónicos dentro de una pastilla minúscula de silicio. En las décadas de 1950 y 1960, las memorias electrónicas estaban formadas por hileras de núcleos o anillos de un material de ferrita, de un milímetro de diámetro o menos, ensartados por miles en rejillas de alambres. Actualmente, en los nuevos proyectos, las memorias de núcleos de ferrita han sido ampliamente sustituidas por memorias de semiconductor, que proporcionan un acceso mucho más rápido a los datos, un tamaño físico mucho más pequeño y un menor consumo de potencia; todo ello a un coste significativamente más bajo. En los años 1980 pueden entrar en escena nuevas tecnologías de memoria a base de burbujas magnéticas, dispositivos superconductores de unión-túnel y otros con acceso por haces de rayos láser o de electrones. Más adelante describiré algunos de ellos. Pero el hecho es que, al ser las memorias de semiconductor extremadamente versátiles, muy compatibles con otros dispositivos electrónicos en sistemas tanto pequeños como de gran capacidad y, al mismo tiempo, susceptibles de mejoras adicionales en rendimiento y en coste, se espera que dominen el mercado de las memorias electrónicas por lo menos durante otra década.

El bloque más pequeño de información accesible en un sistema de memoria puede ser un bit único (representado por 0 o 1), un grupo mayor de bits, tal como un octeto o un carácter (usualmente 8 o 9 bits) o una "palabra" (12 a 64 bits dependiendo del sistema particular). La mayor parte de las memorias son de ubicación dirigida, lo que significa que un bit, octeto o palabra deseados tienen una dirección o posición física especificada a la que se asignan.

Las memorias de contenido dirigido, en las cuales el conocimiento parcial de una palabra almacenada es suficiente



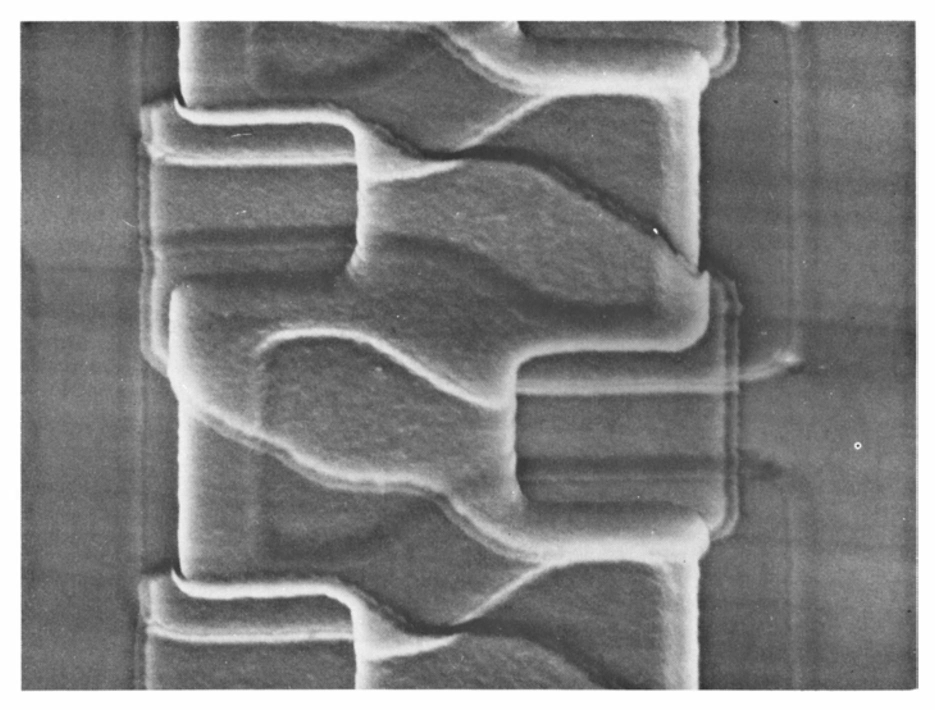
LAS MEMORIAS DE ACCESO ALEATORIO están dispuestas usualmente en conjuntos rectangulares de filas y de columnas. El diagrama de la izquierda muestra un conjunto de ocho por ocho para almacenar 64 bits, almacenándose un bit en la célula que hay en cada intersección. Para especificar una posición determinada de la memoria se necesitan tres dígitos binarios para indicar la fila y otros tres para indicar la columna. En este ejemplo, la dirección de la fila 011 (número binario correspondiente a 3) y la dirección de la columna 101 (número binario correspondiente a 5) especifican la posición 3,5 de la memoria. (Las posiciones comienzan con 0,0 en la parte superior izquierda y terminan con 7,7 en la parte inferior derecha, especificando un número total de 64 posiciones.) A la derecha de la figura se representa el esquema de una célula de memoria de un solo transistor. La información binaria se almacena en forma de carga en un pequeño condensador. Por ejemplo, la carga cero puede representar el número binario 0 y una carga de 500×10^{-15} culombios puede representar el número binario 1. Cuando se activa una de las líneas de selección, o filas, de un conjunto (aquí es la fila 3) se excitan todos los transistores de conmutación conectados a ella. El transistor funciona como un interruptor todo-nada para conectar el condensador de almacenamiento a su línea de datos particular, que corresponde a una columna del conjunto, La activación simultánea de una fila y de una columna identifica la célula seleccionada para la lectura o la escritura (aquí es la célula 3,5). Como el condensador de almacenamiento pierde carga por ambas causas, la lectura y la fuga, debe ser regenerado periódicamente, usualmente una vez cada dos milisegundos. La carga regenerada, que suministra el amplificador de umbral, se devuelve al condensador mediante el cierre del interruptor en la línea de datos. Los transistores realizan las operaciones de conmutación.

para encontrar la palabra completa, serían extremadamente útiles en algunas aplicaciones. Las memorias electrónicas de contenido dirigido nunca se han empleado ampliamente, ya que su coste por bit es mucho más alto que el de las memorias de posición dirigida.

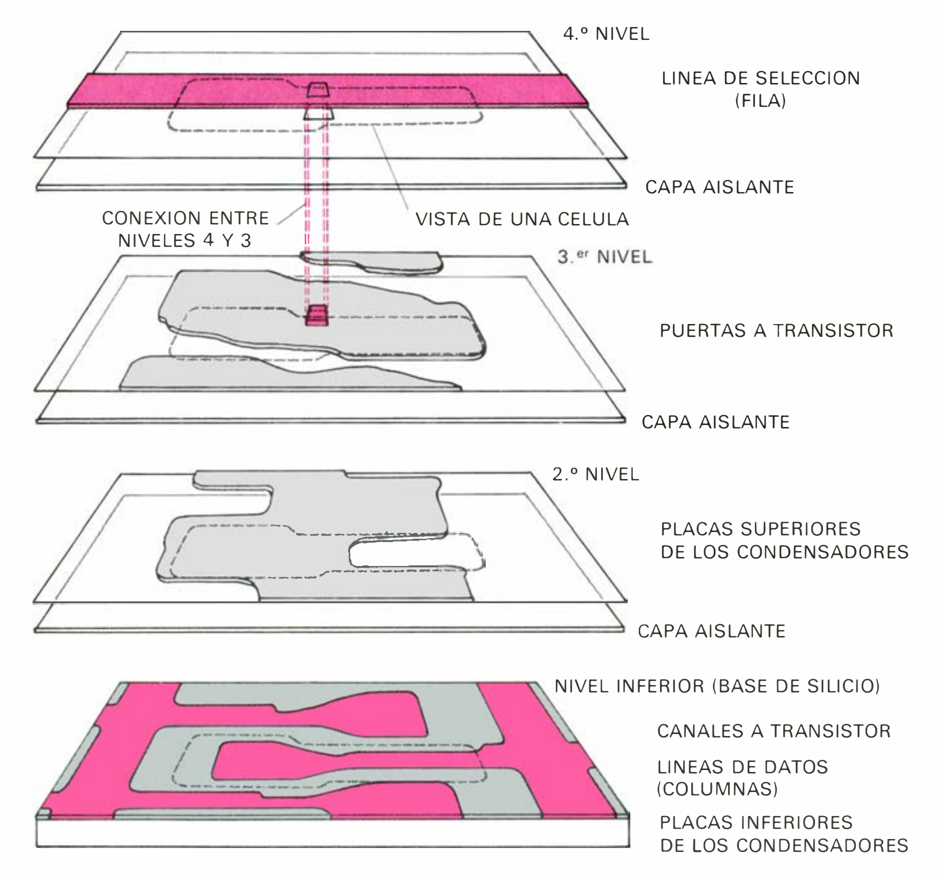
La forma de memoria electrónica más frecuentemente utilizada es la memoria de escritura-lectura de acceso aleatorio, fabricada como una pastilla única de memoria con la técnica de la integración en gran escala, y que es capaz de almacenar un número de bits que puede llegar a 16.384 en un área de menos de medio centímetro de lado. Cierto número de circuitos individuales, cada uno de los cuales almacena un bit binario, se disponen en un conjunto rectangular. El acceso a la ubicación de un bit determinado se logra por medio de una dirección en código binario, presentada en forma de entrada a los decodificadores de dirección; éstos, en cada operación de lectura o de escritura, seleccionan una fila y una columna determinadas. El elemento de memoria situado en la intersección de la fila y de la columna seleccionadas es el único blanco de la lectura o escritura de un bit de información. Una señal de control de lectura-escritura determina cuál de las dos operaciones hay que llevar a cabo. El conjunto de la memoria puede proyectarse con una línea única de entrada-salida para la transferencia de datos o con varias líneas paralelas para la entrada o salida simultáneas de cuatro, de ocho o de más bits.

Tal como señala William G. Oldham en el artículo precedente (véase "La fabricación de los circuitos microelectrónicos", en este mismo número), el coste de un circuito microelectrónico es proporcional al área que ocupa en una oblea de silicio. De aquí que se haya dedicado mucho ingenio al desarrollo de circuitos de memoria simples que ocupan un área muy pequeña. El circuito individual, o "célula", que almacena un bit, es un elemento crítico del proyecto, porque debe repetirse miles de veces. Los decodificadores de dirección y otros circuitos complementarios pueden ser más complejos, porque se emplean una sola vez por fila o por columna e incluso con menor frecuencia.

Hoy en día, los dispositivos de memoria más económicos de lectura-escritura y de acceso aleatorio utilizan ampliamente una célula de almacenamiento sencilla de un solo transistor. La información se almacena en forma de carga eléctrica en un pequeño condensador. El valor de la capacitancia es del orden de 50 femtofaradios (50×10^{-15} faradios). Para la representación de la información binaria se necesitan dos niveles distintos de carga almacenada. El número binario 0 puede representarse mediante una carga cero y el binario 1 por una carga de 500 femtoculombios (equivalente a 10 voltios en el



DOS CELULAS DE ALMACENAMIENTO RAM aparecen en esta microfotografía electrónica de barrido realizada por el autor. Cada célula incluye un transistor y un condensador. Para una mayor claridad, la fotografía, perteneciente a un circuito fabricado por la compañía Intel, se tomó después de haberse realizado sólo tres de las cinco etapas principales de fabricación. La cuarta etapa forma una serie de pequeñas aberturas, una para cada transistor, en la capa aislante que cubre el tercer nivel. La última etapa crea la configuración metálica necesaria para las líneas de selección (filas) en el conjunto final de la memoria. Tal como puede verse con toda claridad en la vista inferior, el metal entra en contacto con las puertas del transistor en el tercer nivel.



LA FABRICACION DE UNA CELULA RAM empieza con la creación de estructuras de usos múltiples *(color)* en el nivel inferior mediante la impurificación con arsénico de ciertas áreas de la capa inferior, que las convierte en un silicio tipo *n*. El área no impurificada se transforma químicamente en nitruro de silicio, un aislador *(gris)*. El segundo y el tercer niveles de la célula se forman por la deposición de silicio policristalino; el cuarto, por la deposición de aluminio. Capas delgadas de dióxido de silicio se depositan entre niveles. La célula mide 15 por 30 micrometros.

condensador de almacenamiento). Si bien esto puede parecer una cantidad minúscula de carga, es realmente la carga de tres millones de electrones. En último extremo, debería poder conseguirse un almacenamiento binario confiable con una carga mil veces más pequeña.

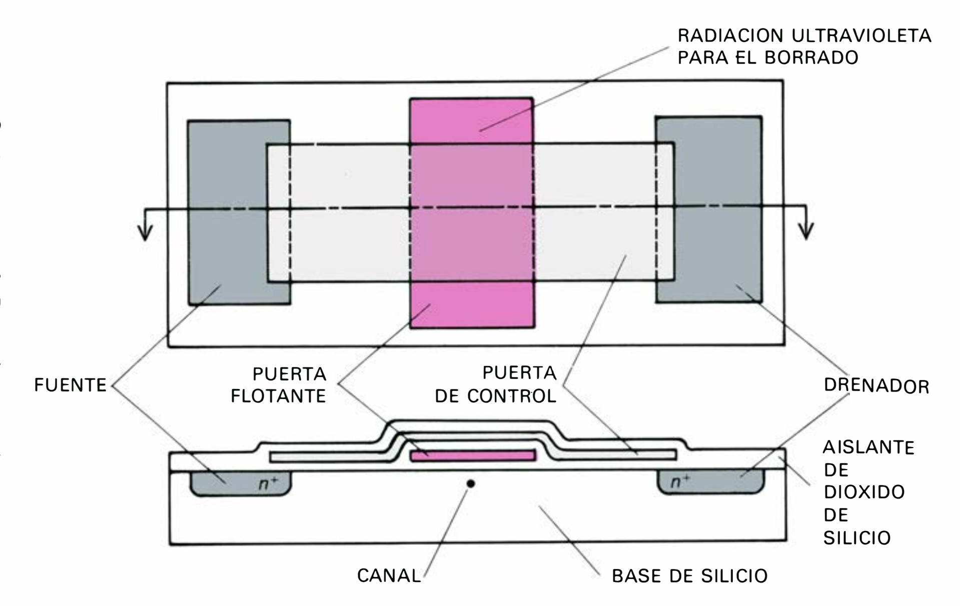
En la célula, el transistor funciona solamente como un interruptor todo-nada que conecta el condensador de almacenamiento a la línea de datos. La línea de datos es compartida por muchas células idénticas, pero sólo una de ellas está conectada cada vez. La línea de datos corresponde a una columna en el conjunto de la memoria. Cada vez sólo se activa una de las muchas líneas de selección correspondientes a las filas que, dentro del conjunto, forman ángulos rectos con las columnas. Cuando se activa una línea de selección, se excitan todos los interruptores de transistor conectados a la misma, pero sólo uno de ellos está en conexión con la línea de datos simultáneamente activada.

La célula de almacenamiento de condensador puede perder la información almacenada de dos formas. En primer lugar, el propio condensador tiene una corriente de fuga significativa, que conduce a la pérdida de la carga almacenada en un tiempo muy corto, de unos pocos milisegundos. En segundo lugar, cuando la célula resulta seleccionada para una operación de lectura, la carga que almacena está compartida entre el condensador de la célula y la capacidad elevada de la línea de datos. En proyectos típicos, la capacitancia de la línea de datos es de 10 a 20 veces mayor que la capacitancia de la célula, puesto que la línea de datos está conectada a muchas células. Como resultado de ello, la tensión que representa la información almacenada resulta atenuada en un factor de 10 a 20. Por tanto, cuando la información almacenada deba retenerse, habrá de ser regenerada después de cada operación de lectura.

También es necesario regenerar periódicamente la carga almacenada, usualmente cada dos milisegundos, al objeto de compensar las fugas. Las funciones de regeneración necesarias se llevan a cabo mediante un amplificador de umbral. Un interruptor de muestreo toma una muestra de la señal en la línea de datos en el momento oportuno: inmediatamente después de que una señal en la línea de selección haya conectado una célula de almacenamiento seleccionada a la línea de datos. La señal de muestreo se compara con un nivel de umbral seleccionado, equidistante entre los niveles bina-

rios nominales de 0 y 1. Seguidamente, el amplificador regenera el nivel binario.

Sólo existe un amplificador de umbral en cada línea de datos, y es de tiempo compartido entre todas las células de almacenamiento conectadas a aquella línea, cuyo número suele ser de 64 o 128 en los diseños actuales. Si cada célula requiere un ciclo de regeneración cada dos milisegundos y un ciclo único ocupa 400 nanosegundos, los ciclos de regeneración para 64 células ocuparán 25,6 microsegundos de cada intervalo de dos milisegundos y los ciclos para 128 células ocuparán el doble de este tiempo, es decir, 51,2 microsegundos. La necesidad de las operaciones de "renovar" o regenerar los datos añade más complicaciones al proyecto de los sistemas de memoria. (Si bien, tal como explicaré más adelante, existen formas para evitar las operaciones de renovación, hacen aumentar el coste del dispositivo.) La función de "escritura" se lleva a cabo por interruptores de accionamiento que están bajo el control de la señal de entrada lectura-escritura.



CELULA DE MEMORIA SOLO DE LECTURA DE BORDADO OPTICO, que puede verse en planta y en sección transversal. La puerta flotante, que no está conectada eléctricamente a ningún circuito, mantiene la información de un dígito binario en forma de una carga almacenada que modifica las características eléctricas del dispositivo. La información se escribe (almacena) en la puerta flotante mediante la aplicación de unos 25 voltios a la puerta de control y al electrodo de drenaje, mientras el surtidor y el substrato están puestos a tierra. Los campos eléctricos elevados que resultan en el canal aceleran los electrones a una velocidad considerable. Una pequeña parte de los electrones es capaz de cruzar el delgado aislador y queda atrapada en la puerta flotante. La carga de la puerta no se pierde durante las operaciones de "lectura" ni cuando se interrumpe la alimentación externa. Sin embargo, la carga almacenada puede borrarse, exponiendo la célula a radiación ultravioleta, que pasa temporalmente el dióxido de silicio a un estado de conducción suficiente como para que la carga se disipe. Las células de la memoria se borran a la vez.

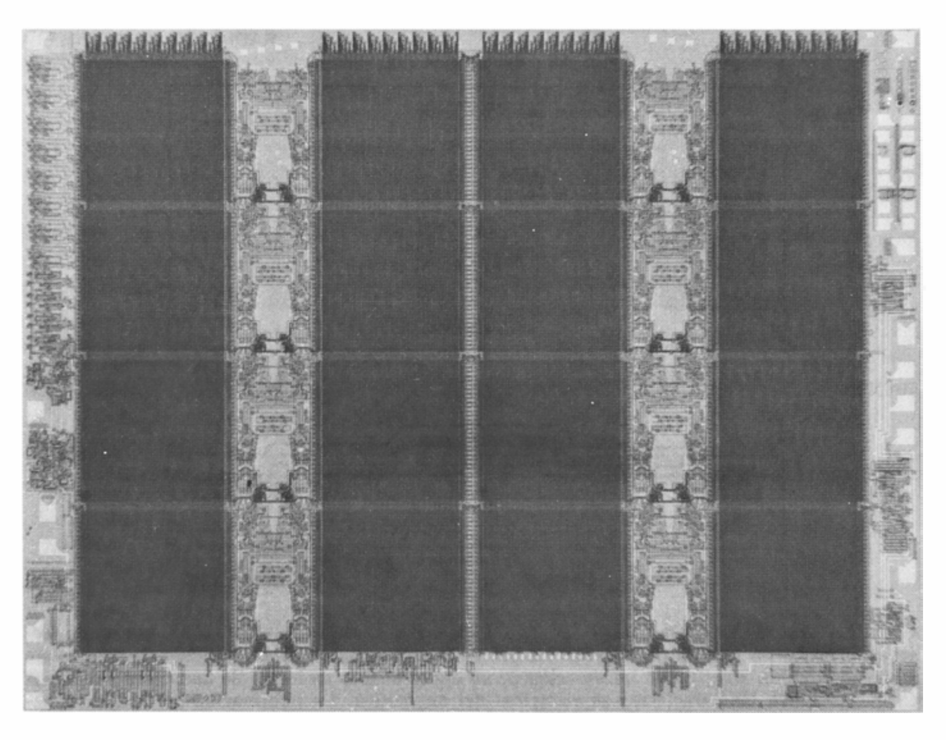
Se necesitan otros circuitos para constituir una pastilla de memoria de semiconductores que sea útil. Los más importantes son los decodificadores de dirección, que seleccionan una fila y una columna únicas para cada entrada de dirección única. Los decodificadores de dirección pueden confeccionarse con puertas lógicas que llevan a cabo la función lógica "y". Existe todavía un gran campo para la inventiva en la obtención de una solución de compromiso entre las demandas conflictivas que son la operación a alta velocidad, el bajo consumo de potencia y la mínima área de pastilla (coste más bajo). Una hábil innovación, presentada en algunos proyectos recientes, consiste en la utilización doble sucesiva de las conexiones de entrada de dirección a un componente, en cada ciclo de memoria. En primer lugar, se aplica la dirección de la fila; a continuación, hacia la mitad del ciclo se presenta la dirección de la columna. La ventaja que ofrece el sistema es que el número de conexiones a la red externa se ve reducido a la mitad. El número de conexiones externas (el número de "patillas") es un factor significativo que afecta al coste, la fiabilidad y el tamaño de los componentes microelectrónicos, de tal modo que, sin duda, su reducción a la mitad es muy ventajosa. Otros circuitos periféricos, incluidos usualmente en la pastilla de memoria microelectrónica, son los separadores de datos (que transforman los niveles de las señales de datos que entran y salen de la pastilla) y generadores de tiempo para la secuencia de las funciones.

La pastilla de memoria de semiconductor con un único transistor en cada célula de almacenamiento es sólo uno de los muchos proyectos existentes. Las células de almacenamiento con dos, tres, cuatro, seis y más transistores ofrecen características adicionales útiles, pero tienen el inconveniente del aumento del área del silicio y, por tanto, del coste. Por ejemplo, en la tecnología de los semiconductores de óxido metálico (MOS), las células de almacenamiento con tres transistores amplifican la señal almacenada, reduciendo mucho el problema de detectar la señal de "lectura" en las líneas de datos. Sin embargo, siguen necesitando renovar la señal almacenada cada dos milisegundos. La renovación continua, dentro de cada célula de almacenamiento, puede alcanzarse añadiendo de uno a tres transistores por célula, pero tal diseño, aparte de ocupar más área de silicio, exige el consumo continuo de potencia en cada célula.

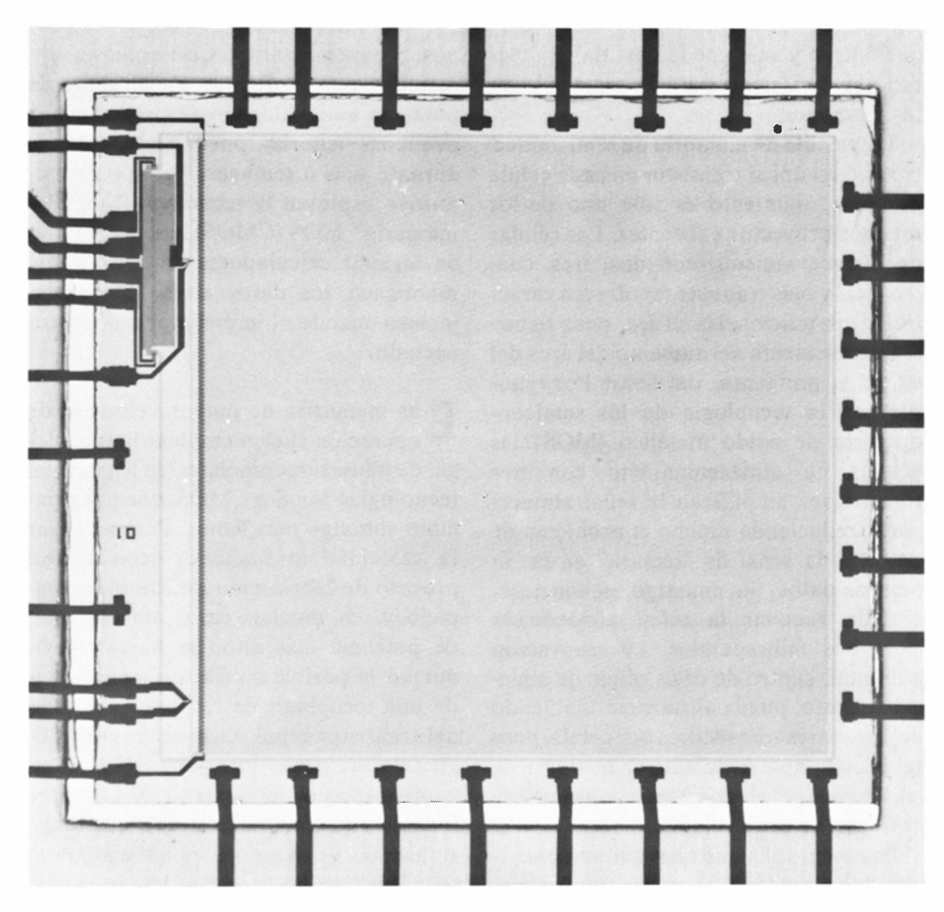
Las memorias que no requieren operaciones de renovación se denominan memorias estáticas. A pesar de su más alto coste por bit de almacenamiento, tienden a usarse preferentemente en los pequeños sistemas de memoria, porque sólo necesitan un mínimo de circuitos externos complementarios. Con un coste adicional, puede reducirse su consumo de potencia a un valor tan despreciable que pequeñas baterías pueden alimentarlas durante días o semanas. Tal tipo de memorias explotan la tecnología "complementaria" MOS (CMOS), y se encuentran en algunas calculadoras de bolsillo que mantienen los datos en su programa, incluso cuando el interruptor está desconectado.

Las memorias de mayor velocidad de operación suelen emplear la tecnología de transistores bipolares, en lugar de las tecnologías MOS o CMOS, que por principio son algo más lentas. Para alcanzar la velocidad más alta es necesario un proceso de fabricación mucho más complejo y, en muchos casos, un consumo de potencia más alto. Se ha estudiado mucho la posible aplicación a memorias de una tecnología de reciente desarrollo del transistor bipolar, conocida como lógica de inyección integrada (I^2L) o lógica de transistor intercalado (MTL), pero todavía no se ha demostrado que tenga suficientes ventajas, ya sea en rendimiento o bien en coste, sobre las memorias basadas en las tecnologías más antiguas de los transistores bipolares.

Algunas aplicaciones requieren memorias de acceso aleatorio que contengan información que esté almacenada de modo



DISPOSITIVO DE CARGA ACOPLADA (CCD), fabricado por la compañía Fairchild Camera and Instrument, que proporciona un acceso en serie a 65.536 bits de memoria. El número total de bits está dividido entre 16 conjuntos cuadrados, manteniendo cada uno 4096 bits. Como los bits circulan continua y sincrónicamente en todos los conjuntos, el tiempo necesario para acceder a cualquiera de los 65.536 bits viene fijado por el tiempo de circulación de los 4096 bits en un conjunto cualquiera. El tiempo medio de acceso es de 0,5 milisegundos. La pastilla mide 4,4 por 5,8 mm.



MEMORIA DE BURBUJA MAGNETICA, fabricada por Texas Instruments Incorporated, proporciona un acceso en serie a 100.637 bits con un tiempo medio de acceso de cuatro milisegundos. Las burbujas magnéticas son dominios, o islas, de polarización magnética embebidos en películas magnéticas delgadas de polarización opuesta. Las burbujas pueden circular a lo largo de las vías prescritas a alta velocidad y no desaparecen cuando se desconecta la alimentación. La densidad en los dispositivos es de unos 150.000 bits por centímetro cuadrado: algo menos que en los CCD.

permanente o que se altere pocas veces. Por ejemplo, es normal que los programas de control de las calculadoras de bolsillo estén almacenados de modo permanente. Tal almacenamiento lo proporcionan las memorias sólo de lectura (ROM). La información se sitúa en la matriz de almacenamiento cuando se fabrica la pastilla. Para obtener una memoria sólo de lectura en una célula de memoria de un transistor puede sustituirse el condensador de almacenamiento por un circuito abierto o bien por una conexión a tierra, representándose así uno u otro de los dos estados binarios. Como alternativa para formar este conjunto de conexiones en la fabricación inicial de la memoria, las células pueden fabricarse con una pequeña conexión fusible desde el transistor a tierra, en lugar del condensador de almacenamiento. Para ubicar luego en la matriz una configuración de información dada, basta con aplicar un conjunto de señales eléctricas lo suficientemente intenso como para destruir las conexiones no deseadas. Como es natural, una vez que una conexión fusible ha desaparecido no se puede rehacer. Estas memorias de conexión fusible son uno de los tipos de memorias programables sólo de lectura (PROM). La información almacenada en ROM y PROM es no volátil, es decir, no se pierde cuando se elimina la alimentación externa.

Otra variación de la memoria sólo de lectura es la memoria principalmente de lectura, que es deseable cuando, necesitándose un almacenamiento no volátil, las operaciones de lectura son mucho más frecuentes que las de escritura. Las memorias principalmente de lectura adoptan dos disposiciones. La más común es la memoria sólo de lectura de borrado óptico. Esta memoria es leída y escrita por medios enteramente electrónicos, pero exige que, antes de efectuarse una operación de escritura, se borren todas las células de almacenamiento, colocándolas en el mismo estado inicial por medio de la exposición de la pastilla completa a una radiación ultravioleta.

En la forma reciente de estas pastillas de memoria, cada célula contiene un transistor MOS con dos electrodos puerta, uno encima del otro. La puerta más baja es del tipo "flotante" y está rodeada totalmente por dióxido de silicio, que es un aislador excelente, y no está conectada eléctricamente a ningún circuito. La tensión de umbral necesaria para excitar el transistor por el electrodo de puerta superior está controlada por la carga de la puerta flotante. La puerta flo-

tante aislada puede cargarse con electrones aplicando una tensión bastante alta (alrededor de 25 voltios) en la puerta y en uno de los electrodos de drenaje, mientras se mantiene a una tensión baja el material de tipo *p* del substrato del dispositivo. Este sistema proporciona a algunos electrones la suficiente energía (alrededor de cuatro electrónvoltios) como para cruzar la barrera aislante de dióxido de silicio y cargar la puerta flotante. Procediendo así en forma selectiva en un conjunto de células, es posible almacenar el conjunto de información deseado.

La alta calidad del aislador de dióxido de silicio permite que la puerta flotante permanezca cargada con seguridad durante años. Para descargar la puerta flotante, se expone la pastilla completa de memoria a la radiación ultravioleta, lo que hace que el dióxido de silicio se vuelva lo suficientemente conductor como para permitir que la carga de la puerta fluya. Debido al tamaño minúsculo de las células individuales de almacenamiento, no resulta práctico borrarlas selectivamente. Se prefiere borrar simultáneamente todas las células de memoria y cargarlas después con una nueva información.

Otra forma de memoria principalmente de lectura es la memoria eléctricamente alterable sólo de lectura (EAROM, del inglés "electrically alterable read-only memory"), que puede modificarse sin necesidad de un borrado total. La estructura de su célula es bastante similar a la de la célula de borrado óptico, salvo que la función de la puerta flotante se lleva a cabo mediante una interfase entre dos materiales dieléctricos, que suelen ser el dióxido de silicio y el nitruro de silicio; la célula puede cargarse y descargarse selectivamente mediante señales aplicadas a un único electrodo de puerta superpuesto. Estas memorias principalmente de lectura se denominan dispositivos semiconductores de óxido de nitruro metálico (MNOS, del inglés "metal-nitride-oxide-semiconductor"). Una variante de la memoria principalmente de lectura se construye a partir de ciertos materiales semiconductores amorfos, o vidrios semiconductores, que mantienen una estructura de cargas locales hasta que un fuerte impulso de corriente borra las cargas, permitiendo que después se imprima eléctricamente una nueva distribución. Si bien los aparatos EAROM tienen una gran utilidad potencial, no disponen todavía de la combinación de bajo coste, escritura rápida, período largo de retención y fiabilidad

después de muchos borrados que se necesita para la mayor parte de las aplicaciones, por lo que su utilización actual es todavía limitada.

Todas las memorias electrónicas comentadas están proyectadas para el acceso aleatorio, es decir, el tiempo requerido para una operación de lectura o de escritura es independiente de la posición física que ocupe la célula dentro del conjunto de almacenamiento. Los tiempos de acceso de las memorias de acceso aleatorio son del orden de 0,1 microsegundos a un microsegundo. Prescindiendo del acceso aleatorio pueden lograrse importantes economías en el diseño de las memorias. Un esquema alternativo de importancia creciente es el acceso en serie, en el cual los bits almacenados circulan como si estuviesen en una tubería cerrada. Ejemplos importantes de dispositivos de memoria electrónica organizados de esta forma son los dispositivos de carga acoplada (CCD, del inglés "charge-coupled devices") y los de burbuja magnética. Cada bit almacenado se transfiere secuencialmente a través de 64 o más posiciones de almacenamiento desde el momento en que se escribe en la memoria hasta el primer instante en que está disponible para la lectura. La velocidad con la que los bits de una memoria CCD pasan de una célula de almacenamiento a la siguiente es aproximadamente la misma que la duración del ciclo requerido para una memoria de acceso aleatorio. Por consiguiente, el tiempo de acceso más largo, en una memoria en serie con 64 posiciones de almacenamiento, es 64 veces la duración del ciclo de una memoria de acceso aleatorio. Existen muchas aplicaciones en las que el acceso en serie resulta enteramente satisfactorio. Por ejemplo, una memoria que se emplee para renovar la información presentada en un terminal de video convencional, que se explora punto por punto en forma lineal repetitiva, no requiere una memoria con acceso aleatorio.

Existen varias razones por las que las memorias CCD pueden diseñarse de modo tal que su área total por bit sea más pequeña (y, por tanto, su coste sea potencialmente más bajo) que las memorias de semiconductor de acceso aleatorio. Primeramente, cuando la tecnología de proceso se optimiza para usos CCD, el área requerida para una célula de almacenamiento individual es algo más pequeña que la necesaria para una célula RAM. En segundo lugar, la cantidad de decodificación de dirección necesaria en una memoria en serie es por naturaleza menor que en una RAM, ya que no hace falta la decodificación para seleccionar posiciones individuales. En tercer lugar, la carga que representa la información almacenada se transporta y se mantiene siempre en nodos o intersecciones minúsculas del circuito, que poseen muy poca capacitancia. Como resultado, la señal de tensión no se atenúa en una CCD como lo es en una memoria de acceso aleatorio, donde la señal es compartida entre la célula de almacenamiento y la línea de datos. Por tanto, los circuitos amplificadores y de renovación de la memoria CCD pueden ser algo más sencillos. El resultado neto de estas simplificaciones es que el área total de silicio por bit de los componentes de la memoria completa es aproximadamente de dos a tres veces menor en una memoria CCD que en una memoria de acceso aleatorio. Uno de los últimos proyectos realizados para una memoria de acceso en serie CCD puede almacenar 65.536 bits en una pastilla que mide sólo 3,5 por 5 milímetros.

La otra forma importante de memoria microelectrónica de acceso en serie se basa en la movilidad de las burbujas magnéticas o dominios microscópicos de la polarización magnética en una película magnética delgada de ortoferrita o granate. En presencia de un campo magnético fijo de intensidad adecuada, con sus líneas de fuerza perpendiculares al plano de la película, los dominios de una polarización (por ejemplo, hacia arriba) son estables dentro de una amplia área de la polarización opuesta (hacia abajo). Los dominios pueden moverse en el plano de la película mediante la aplicación de campos magnéticos más débiles que forman ángulos rectos con el

DIAGRAMA DE UNA MEMORIA CCD, que describe la forma de circulación de 4096 bits almacenados en cada conjunto del dispositivo de memoria de 16 conjuntos y 65.536 bits. Tal como ocurre en la mayor parte de los dispositivos de memoria de acceso aleatorio, es necesario regenerar o renovar continuamente la configuración de los bits almacenados. El ciclo de regeneración establece también el tiempo de acceso para la lectura y la escritura. En este caso, los 64 bits de cada una de las 64 columnas se desplazan sincrónicamente hacia abajo y hacia el regenerador, a lo largo de líneas paralelas con una velocidad de unos 80.000 bits por segundo. Los 4096 bits (64 × 64) fluyen en serie fuera del regenerador a una velocidad 64 veces más alta (es decir, alrededor de cinco millones de bits por segundo) y entran de nuevo por la parte superior del conjunto, distribuidos a lo largo de 64 líneas paralelas a sus columnas originales. Como los 16 conjuntos operan de una forma similar, el dispositivo de 65.536 bits no puede distinguirse externamente de un registro desplazador en serie de 4096 bits que trabaje a cinco millones de ciclos por segundo.

campo principal. Cuando los dominios están adecuadamente organizados, pueden crear una memoria de burbuja magnética.

Si se emplea un pequeño dominio con polarización hacia arriba para designar el número binario 1, es claro que la ausencia de tal dominio puede considerarse como el número binario 0. Los dominios pueden formarse o eliminarse con lazos electromagnéticos de una sola espira dispuestos sobre la superficie de la película magnética. Para crear variaciones locales en el campo magnético y situar y separar así las burbujas individuales, se utilizan conjuntos de electrodos en forma de *chevron*, de tes o de discos fabricados con otro material magnético. Puede emplearse un campo magnético rotativo situado en el plano de la película magnética para correr los dominios a lo largo de una línea fija de *chevrons* (o de alguna otra disposición).

Una característica capital de las memorias de burbuja es que retienen la información almacenada cuando se interrumpe la alimentación externa, lo cual constituye una propiedad valiosa que se aprovecha en la mayoría de las aplicaciones actuales. Para preservar la polarización de las burbujas se utiliza un imán permanente que mantiene fijo el campo magnético perpendicular. Es necesario emplear circuitos electrónicos para excitar los lazos de formación de burbujas y los de mando y para amplificar la pequeña señal de salida que se obtiene mediante inducción electromagnética o por un efecto magnetorresistivo cuando una burbuja pasa a través de un dispositivo sensor. Estos circuitos se alimentan durante las operaciones de lectura y de escritura.

Las memorias de burbuja están organizadas intrínsecamente en serie, por lo que el tiempo de acceso depende del número de posiciones de almacenamiento que existan en un trayecto y de la velocidad máxima de desplazamiento. En los aparatos actuales, los trayectos seriales varían en longitud desde unas 10 posiciones hasta 1000 o más; las velocidades de desplazamiento varían desde una fracción de microsegundo hasta varios microsegundos. Análogamente a las memorias CCD, las memorias de burbuja no pueden competir en velocidad con las memorias electrónicas de acceso aleatorio. Su aplicación potencial más prometedora reside en la sustitución de las memorias de cinta y de disco de capacidad comprendida entre un millón y 10 millones de bits. En tales aplicaciones, las memorias de semiconductores pueden ser desplazadas porque su almacenamiento de información es volátil. Actualmente, el coste por bit de las memorias de burbuja, si se fabrican en grandes cantidades, es aproximadamente comparable al de las memorias de superficie móvil para capacidades de almacenamiento de hasta unos 10 millones de bits. El coste de las memorias de superficie móvil posee todavía una importante ventaja cuando se trata de almacenar mayores cantidades de información.

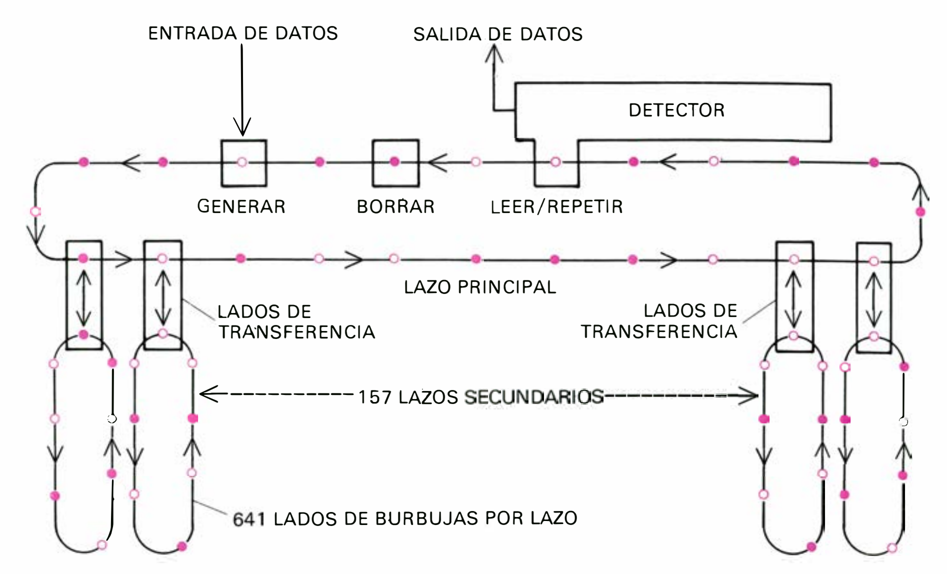


DIAGRAMA DE MEMORIA DE BURBUJA MAGNETICA, que muestra el esquema de circulación en una memoria de 100.637 bits. El lazo principal mantiene un bloque único de datos, formado por los unos (1) y los ceros (0) (burbujas o no burbujas) que se están escribiendo, leyendo, reproduciendo o borrando en la memoria. En este dispositivo concreto, el bloque de datos contiene 157 bits. En el ciclo de escritura, los 157 bits entran en primer lugar en el lazo principal, donde a una señal dada son transferidos simultáneamente hacia los 157 lazos secundarios, con un bit por lazo. A su vez, cada lazo secundario proporciona espacios para 641 burbujas. Por consiguiente, la capacidad total del aparato es de 157 × 641, o sea, 100.637 bits. En el ciclo de lectura, y a una señal dada, los 157 bits se transfieren simultáneamente desde los lazos secundarios al lazo principal.

Una variación prometedora de la memoria de burbuja es la red de registro de burbujas o circuito de burbuja sin estructura, que elimina la necesidad de uno o dos electrodos de superficie mecanizada por bit de información almacenada. Las burbujas se almacenan muy cerca unas de otras; la distinción entre estados binarios viene determinada por cambios en la magnetización dentro de la pared de un dominio simple más que por la presencia o ausencia de un dominio. El desarrollo adecuado de las memorias de red de burbujas vencería las limitaciones impuestas por las técnicas empleadas para definir las estructuras, que es lo que determina actualmente la máxima densidad de almacenamiento.

El dispositivo que quizás ofrece mayores posibilidades para el perfeccionamiento de las memorias microelectrónicas, y que puede llegar a hacer posible la creación de una nueva generación de computadores superrápidos, es la unión túnel superconductora. Tales uniones son dispositivos de conmutación de terminales múltiples que, en respuesta a un pequeño cambio en el campo magnético local, gobiernan la transición entre un estado de filtración superconductora y un estado de filtración normal. La "filtración" se refiere a la capacidad, que, bajo ciertas condiciones, poseen los electrones para penetrar a través de las barreras de energía, y de la que normalmente carecen al no disponer de la suficiente energía para superarlas. Para que la unión túnel pase al estado de superconducción, hay que enfriar el dispositivo a temperaturas criogénicas mediante la inmersión en helio líquido.

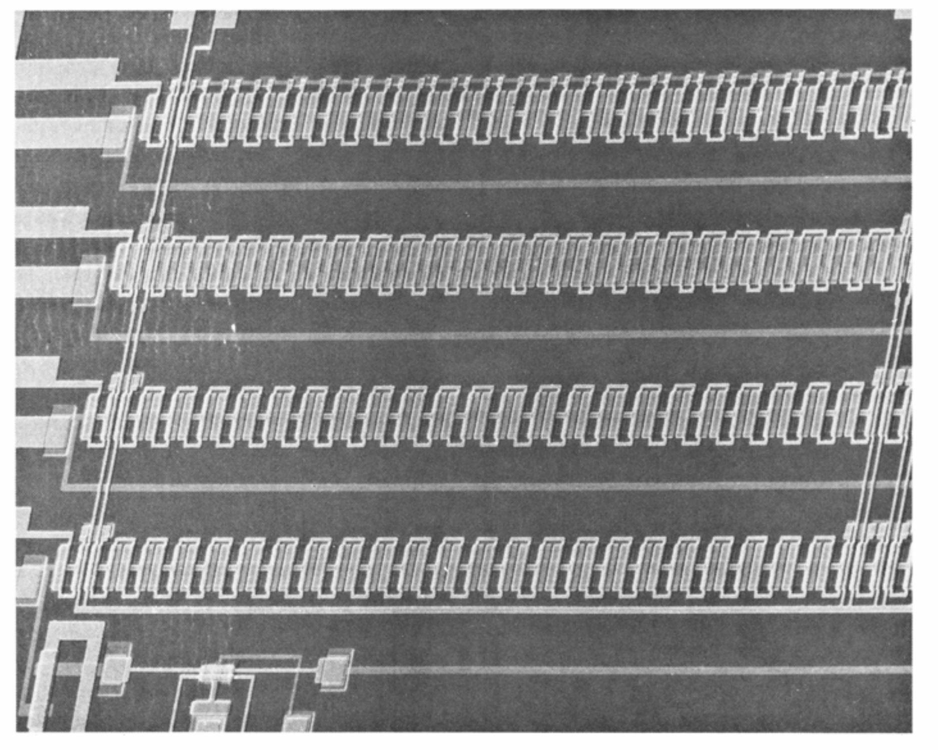
Como las uniones túnel trabajan a temperaturas cercanas al cero absoluto y a tensiones muy bajas, características del estado de superconducción, son capaces de llevar a cabo operaciones de conmutación de uno o dos órdenes de magnitud más rápidas que las que pueden realizar los circuitos a semiconductores y consumen una potencia de alimentación dos o tres órdenes de magnitud menor que éstos. La reducción en el consumo de potencia permite empaquetar los circuitos en forma muy densa sin que se creen problemas de disipación de calor. También se reducen los tiempos de propagación de la señal en un gran sistema, ya que las líneas de interconexión pueden ser más cortas. Estas consideracio-

nes permiten prever el desarrollo de memorias crioelectrónicas con densidades de componentes extremadamente altas y que trabajen a velocidades de 10 a 100 veces mayores que las memorias electrónicas actuales más rápidas.

Existe todavía un número de problemas difíciles de resolver antes de que se implanten las memorias y los computadores crioelectrónicos. La fabricación de las uniones túnel precisa de unas películas aislantes perfectas (usualmente óxidos amorfos que crecen *in situ*) con un espesor aproximado de un doscientosavo de la longitud de onda de la luz azul. Es necesario crear nuevas técnicas de empaquetado e interconexión para alcanzar la alta densidad requerida para lograr retardos mínimos en la interconexión. Conviene estudiar debidamente las tensiones mecánicas generadas por los distintos coeficientes de dilatación térmica para evitar daños cuando la temperatura baja desde la ambiente hasta casi el cero absoluto. Tales consideraciones sugieren que serán necesarios algunos años de trabajo antes de que los sistemas crioelectrónicos sean utilizables.

Hasta aquí he descrito sobre todo las características funcionales y de alta velocidad de las distintas tecnologías de memoria. Al llegar a este punto desearía comentar brevemente los temas que pueden ser los más importantes en la selección de la tecnología de memoria: el precio por bit de la capacidad de almacenamiento y la fiabilidad. Por último, habré de examinar las perspectivas de ulteriores desarrollos que presentan la tecnología de las memorias de superficie móvil y la de las memorias completamente electrónicas.

El precio por bit de un sistema de memoria está correlacionado aproximadamente con la complejidad de la estructura o del sistema, tomando como base un bit. Por ejemplo, las memorias electrónicas de mayor velocidad requieren varios transistores, resistencias y líneas de interconexión para cada célula. Por el contrario, las memorias electrónicas más sencillas necesitan sólo un transistor y dos líneas de interconexión por bit. La fabricación de las memorias de semiconductor exige de cuatro a ocho procesos secuenciales de transferencia de diagramas. Las memorias de burbuja son potencialmente más baratas que las memorias de semiconductores, porque sólo requieren de uno a tres procesos de transferencia de diagramas en la fabricación del conjunto de células de almacenamiento.

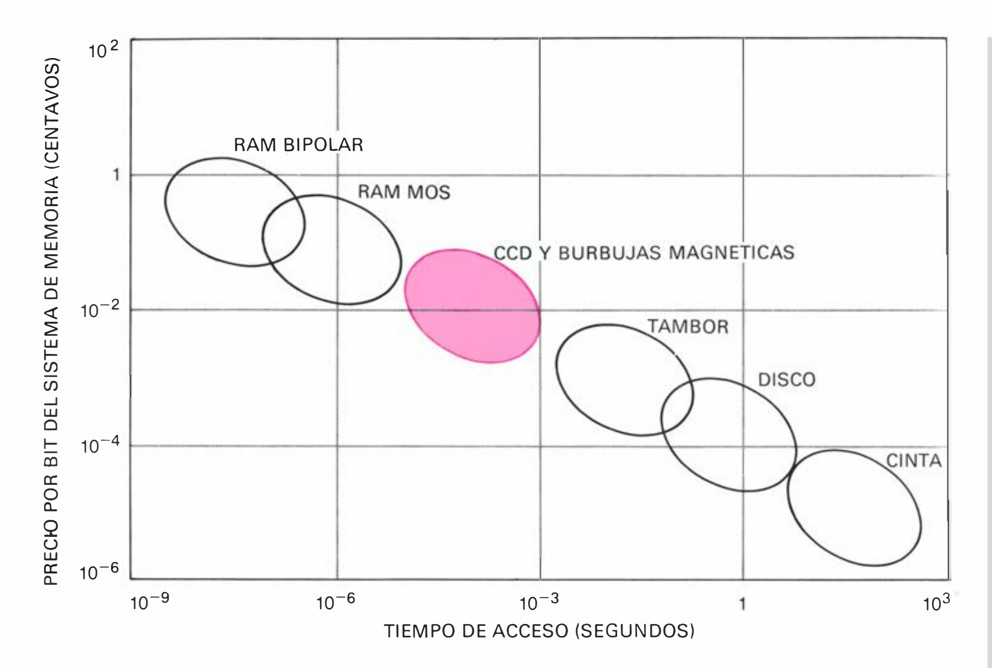


LAS CELULAS DE MEMORIA JOSEPHSON, que trabajan en condiciones cercanas al cero absoluto, pueden almacenar información en forma de cuantos únicos de flujo magnético, llegando, por tanto, a la economía límite de energía alcanzable en un dispositivo magnético de almacenamiento. Este dispositivo experimental, fabricado por el Zurich Research Laboratory de la International Business Machines Corporation (IBM), posee 392 células de memoria Josephson dispuestas en cuatro hileras de 98 células cada una. La pastilla de ensayo, de la que puede verse sólo una sección, fue fabricada en parte para ver si la entrada o la salida de información perturbaría la información que estaba almacenada en otras células distintas. Los resultados que se obtuvieron confirman la posibilidad de almacenar un cuanto único en células que se encontraban muy poco separadas.

Las memorias de superficie móvil son más baratas que las memorias electrónicas porque en ellas no es necesario definir diagramas o estructuras individuales para cada célula de almacenamiento individual. Sin embargo, el precio mínimo de un sistema de memoria de superficie móvil es relativamente alto a causa de la necesidad de fabricar componentes mecánicos de precisión para transportar el medio de almacenamiento magnético. Lo que hace atractivas desde el punto de vista económico a las memorias de red de burbujas es que no requieren ni partes móviles ni definir características en cada célula.

Un método para comparar los precios de las varias formas de memorias es oponerlas en función del tiempo medio de acceso aleatorio. En los sistemas de que ahora se dispone, los intervalos del tiempo de acceso varían desde 10 nanosegundos para las memorias de transistor bipolar hasta 10 segundos o más para las memorias de cinta magnética. El coste correspondiente por bit varía aproximadamente desde un centavo de dólar para las memorias bipolares hasta menos de 10^{-5} centavos para la cinta. Dentro del campo de variación de un millón a uno del precio por bit, el esquema no cambia: es necesario aceptar un aumento de dos órdenes de magnitud en el tiempo de acceso para lograr un ahorro de un orden de magnitud en el precio.

La fiabilidad de los sistemas de memoria es función tanto de problemas de principio como prácticos. Los problemas de principio se refieren a fenómenos tales como la corrosión; han sido bastante estudiados en las tecnologías de memoria bien desarrolladas y no suelen ser la causa de muchos fallos. Los problemas prácticos se encuentran en la fabricación, en el embalaje o en las comprobaciones deficientes y en los errores cometidos en la utilización y en el mantenimiento de los componentes y sistemas. La mayor parte de los fallos que se producen en los sistemas de memoria actuales derivan de estos problemas prácticos. Por ejemplo, en las memorias de semiconductores, el número de fallos que se produce en sus utilizaciones concretas está relacionado más con el número de pastillas empaquetadas separadamente que con el número total de bits almacenado en el sistema.



COMPARACION del tiempo de acceso y el precio de varias tecnologías de memoria. Los aparatos de carga acoplada y las memorias de burbuja magnética llenan un hueco importante entre las memorias ultrarrápidas de acceso aleatorio y las magnéticas más lentas de superficie móvil. En la próxima década, el coste de todos los sistemas de memoria debería bajar en un factor de 10 o más.

Durante muchos años, se ha observado empíricamente que el coste de fabricación de los productos tecnológicamente complejos tiende a disminuir a medida que aumenta la experiencia acumulada en su fabricación. Se cree que existe una relación similar en cuanto al número de fallos y que, para un producto electrónico dado, parece que disminuyen al aumentar la experiencia en su fabricación. Este hecho, cuya representación gráfica se conoce como la curva de aprendizaje o función de progreso, crea una barrera a la introducción de nuevas tecnologías de memoria. En los últimos veinte años, docenas de nuevas tecnologías de memoria han fracasado frente a los perfeccionamientos continuos en las tecnologías de memoria tradicionales. Durante este período, las memorias magnéticas de superficie móvil han mejorado de modo continuo. El único ejemplo que constituye la excepción a la regla es el éxito de las memorias de semiconductor al sustituir a las memorias de núcleo de ferrita como técnica dominante en el almacenamiento electrónico. Muchos observadores piensan que este éxito sólo se debe al hecho de que la tecnología microelectrónica de los semiconductores ya estaba avanzando a lo largo de una pronunciada curva de aprendizaje antes de que se la utilizara para las memorias.

Un cierto número de tecnologías de memoria prometedoras no han alcanzado éxito comercial a pesar de la investigación y el desarrollo intensivos y, en algunos casos, incluso de una producción limitada. Ejemplos de esta clase se encuentran en las memorias de película magnética (tales como las memorias de hilo recubierto y de película planar) y en las memorias de haz electrónico y de haz óptico. A menudo, la inversión requerida para desarrollar una nueva tecnología de memoria hasta llegar a la etapa de fabricación a bajo coste y con alta fiabilidad se ha considerado demasiado grande en relación a los riesgos y resultados potenciales.

Es probable que la próxima década nos ofrezca perfeccionamientos sustanciales en el rendimiento de las memorias de superficie móvil y las electrónicas, con reducciones paralelas en el coste. No existen barreras fundamentales que impidan aumentar en más de cien veces la densidad de almacenamiento de bits en las superficies magnéticas móviles, con un pequeño aumento en el precio del sistema. La introducción prevista de las técnicas de haz electrónico y de rayos X en la fabricación de circuitos microelectrónicos debería posibilitar también el aumento de la densidad de bits de estos dispositivos en un factor de 100, con sólo aumentos pequeños en el precio por componente. Por consiguiente, la esperanza es que, en los próximos diez años, el precio por bit de todas las formas de memoria digital se vea reducido en más de un orden de magnitud.

Microprocesadores

Un microprocesador es la unidad central de proceso de un ordenador contenida en una pastilla. Suele hallarse asociado con otras pastillas dentro de un microcomputador, si bien están surgiendo ya sistemas computadores complejos en una sola pastilla

Hoo-Min D. Toong

El microprocesador es la unidad central aritmética y lógica de un ordenador; junto con sus circuitos asociados, está reducida a una escala tal que permite disponerla sobre una sola pastilla de silicio (a veces varias pastillas), la cual contiene decenas de miles de transistores, resistencias y otros elementos del circuito. Forma parte de la familia de los circuitos integrados a gran escala, que es la que refleja el estado actual de la evolución de un proceso de miniaturización que comenzó con el desarrollo del transistor, a finales de la década de los 40. Una pastilla de microprocesador típica mide medio centímetro de lado. A base de añadirle entre 10 y 80 pastillas para dotarle de temporización, memoria de programa, memoria de acceso aleatorio, circuitos intermedios para entrada y salida de señales y otras funciones auxiliares, se puede construir un sistema computador completo en una placa cuya superficie no sobrepasa el tamaño de esta página. Tal conjunto es un microcomputador, en el que el microprocesador hace de componente principal. Alrededor de 20 empresas norteamericanas fabrican actualmente unos 30 diseños diferentes de pastillas de microprocesador, con una escala de precios entre 10 y 300 dólares. Más de 120 empresas incorporan estas pastillas en sistemas microcomputadores, que se venden a partir de 100 dólares. La cantidad de aplicaciones de los microprocesadores crece día a día en la industria, la banca, la generación y distribución de energía, las telecomunicaciones y en gran número de productos comerciales, que van desde los automóviles hasta los juegos electrónicos.

Como en la unidad central de proceso (CPU, del inglés "central processing unit") de un gran computador, la tarea del microprocesador consiste en recibir datos en forma de cadenas de dígitos binarios (ceros y unos), almacenarlos para su procesamiento posterior, realizar operaciones aritméticas y lógicas con los datos, de acuerdo con instrucciones previamente almacenadas y entregar los resultados al usuario a través de un mecanismo de salida, que puede ser un teletipo, la pantalla de un tubo de rayos catódicos o un trazador de gráficos bidimensionales. El diagrama de bloques de un microprocesador típico mostraría las unidades siguientes: una unidad de decodificación y control (para interpretar las instrucciones del programa almacenado), la unidad aritmética y lógica o ALU (para llevar a cabo operaciones aritméticas y lógicas), registros (que sirven como memoria de datos fácilmente accesibles, para los datos que se manipulan con más frecuencia), un acumulador (registro especial asociado íntimamente con la ALU), separadores o memorias intermedias de direcciones (para suministrar a la memoria de control las direcciones en las que ir a buscar la instrucción siguiente) y separadores de entrada-salida (para leer instrucciones o datos que entren al microprocesador o para enviarlos al exterior).

Los microprocesadores actuales varían en su arquitectura específica según su fabricación, y a veces, según la tecnología de semiconductores adoptada en cada caso particular. Una de las distinciones más significativas consiste en si se encapsulan en una pastilla todos los elementos del microprocesador o bien se reparten entre varias pastillas modulares idénticas, que pueden unirse en paralelo; el número total de pastillas depende de la longitud de la "palabra" que el usuario pretende procesar: cuatro bits (dígitos binarios), ocho bits, 16 bits o más. Tal disposición multipastilla se conoce como organización de bit rebanado ("bit-sliced"). Una característica de las pastillas de bit rebanado ("bit-sliced") hechas con tecnología bipolar es que son "microprogramables": permiten al usuario crear sus conjuntos específicos de instrucciones, lo cual representa una ventaja clara para muchas aplicaciones.

La avalancha de microprocesadores y microcomputadores que llegan al mercado, junto con la elevada tasa de innovaciones, garantiza que cualquier intento de catalogarlos envejecerá instantáneamente. Una introducción más adecuada al mercado de los "micros" consiste en clasificar jerárquicamente los sistemas de acuerdo con su potencia y su función. A lo largo de estas dos dimensiones hay una progresión ascendente bien definida tanto en los componentes físicos como en los componentes lógicos. Por lo que a los componentes físicos se refiere, los niveles son: pastillas, módulos, sistemas en panel, pequeños sistemas computadores, sistemas completos de desarrollo y sistemas multiprocesadores.

Esta jerarquía no es absoluta, ya que la tecnología en su evolución crea pastillas cada vez más potentes, algunas de las cuales pueden solapar dos o tres niveles jerárquicos. Las pastillas se utilizan para construir módulos, los módulos para construir pequeños sistemas computadores y los pequeños computadores para construir sistemas de desarrollo completos. Los sistemas multiprocesadores pueden reunir elementos de las tres últimas categorías, dependiendo de la aplicación y la complejidad.

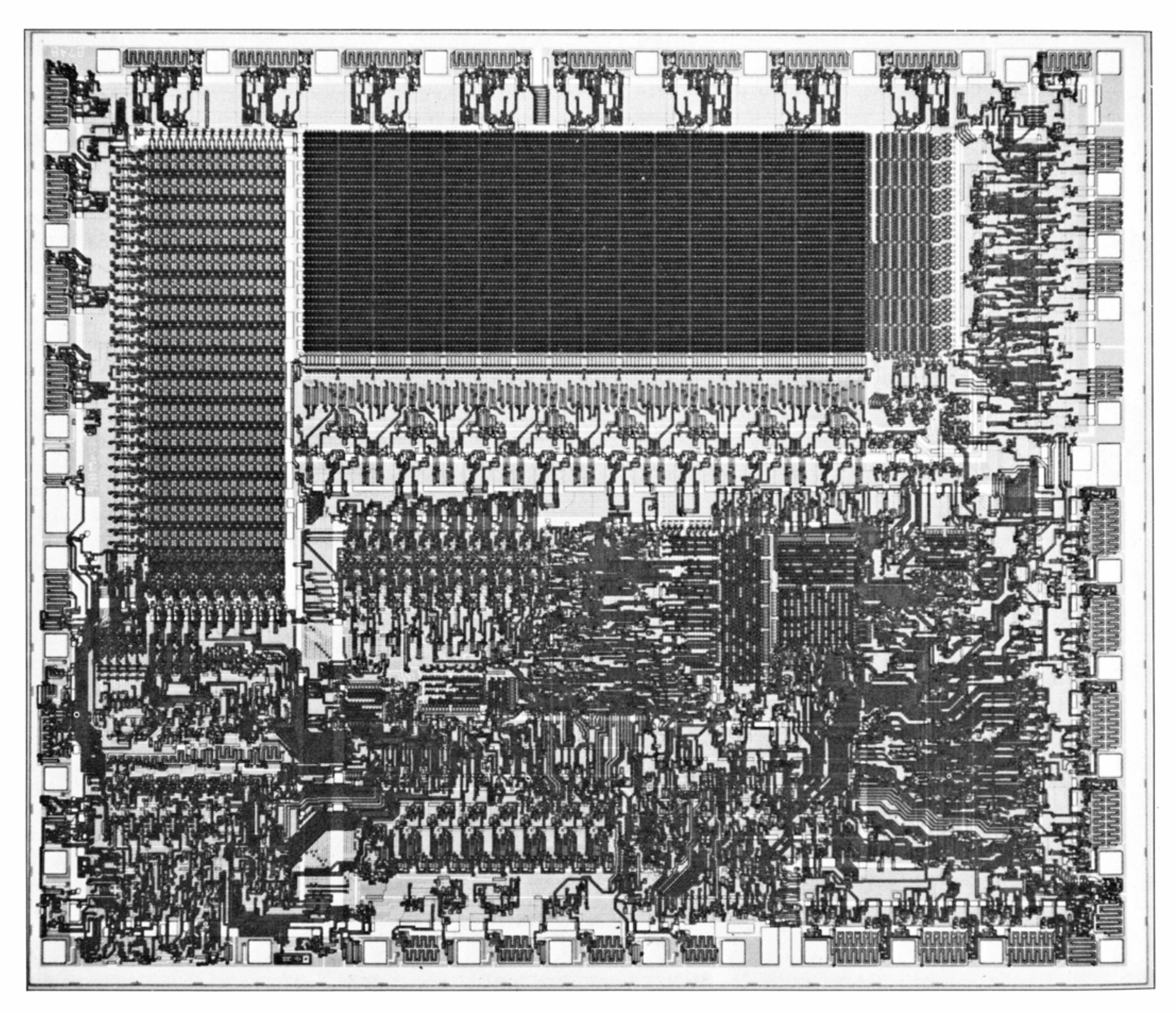
En el primer nivel de la jerarquía se encuentran las pastillas de microprocesador que representan la integración a gran escala de decenas de miles de dispositivos electrónicos individuales: transistores, diodos, resistencias y condensadores.

En este nivel están también pastillas más especializadas: memorias de acceso aleatorio (RAM), memorias de sólo lectura (ROM), memorias de sólo lectura y programables (PROM), circuitos

intermedios de entrada-salida (I/O) y algunas otras. La tecnología de vanguardia trabaja especialmente a nivel de pastilla, proporcionando, por ejemplo, dispositivos RAM con capacidad de almacenamiento cada vez mayor. (Actualmente la RAM más avanzada, y comercialmente disponible, puede almacenar 16.384 bits; dentro de uno o dos años, la capacidad de almacenamiento máxima será cuatro veces mayor.)

Por lo general, los diversos tipos de pastillas se agrupan en familias compatibles con microprocesadores particulares. Las familias incluirán una serie de pastillas RAM, ROM y PROM para crear un sistema de memorización, una serie de pastillas de acoplamiento, capaces de manejar funciones de entrada-salida tanto en serie como en paralelo, y pastillas diversas para aumentar las posibilidades del sistema, tales como operaciones aritméticas a gran velocidad. Se necesitan pastillas de control central que establezcan prioridades y que mantengan el flujo adecuado de las señales a través del complicado laberinto de interconexiones. La compatibilidad entre pastillas y entre familias de pastillas construidas por fabricantes diferentes presentan una amplia variedad. Por ejemplo, los microprocesadores de constructores diferentes no son, por lo general, intercambiables físicamente, mientras que varios tipos de pastillas de memorización suelen serlo.

El segundo nivel de la jerarquía, el módulo y los sistemas en panel, constituyen los sistemas computadores más sencillos, propiamente dichos. Pueden crearse por combinación de un microprocesador con un conjunto limitado de pastillas de memorización (RAM y ROM) y pastillas de entrada-salida. Para comunicarse con un sistema mínimo así, el usuario precisará también un dispositivo sencillo, como lo es un teclado numérico, y un dispositivo capaz de presentar visualmente o de registrar la salida del computador. Estos sistemas, contenidos en una sola placa, son útiles para fines didácticos elementales o pueden servir también como prototipos en panel para sistemas más elaborados. Con una



MICROCOMPUTADOR de una sola pastilla: circuito integrado a gran escala que constituye un sistema completo de control y procesado digital, para aplicaciones generales. El dispositivo combina un microprocesador, que normalmente ocuparía toda una pastilla, con un variado conjunto de funciones suplementarias, como memoria de programa, memoria de datos, acopladores de entrada-salida (I/O) múltiple y circuitos de temporización. El dispositivo que se muestra, un 8748 de la Intel Corporation, mide 5,6 por 6,6 milímetros. El programa se almacena en una EPROM (memoria de sólo lectura, borrable y reprogramable) con un kilooocteto o sea 8192 bits (dígitos binarios) de capacidad. Se borra el programa exponiendo el circuito a una radiación ultravioleta que da lugar a la pérdida de las cargas eléctricas almacenadas en la EPROM; tras ello puede introducirse eléctricamente un programa nuevo. En cantidades de 25 y más unidades, la pastilla 8748 vale 210 dólares.

inversión modesta (menos de 300 dólares normalmente), un principiante puede aprender lo esencial de la programación de microprocesadores. De cualquier modo, debido a la limitación de memoria del sistema, a su falta de elementos de desarrollo del soporte lógico y a su conexión imperfecta con el usuario, incluso un principiante podrá superar en poco tiempo el nivel de un sistema en una sola placa.

En el nivel siguiente de la jerarquía de potencia y función se encuentran los pequeños sistemas computadores, que ya vienen conformados como unidades completas. A diferencia de los módulos en una sola placa, llevan una fuente de alimentación propia, permiten ampliar su capacidad de memoria y tienen provisiones para una serie de módulos de interconexión enchufables. Algunos de los más potentes módulos de desarrollo en una sola placa pueden expandirse con el soporte físico apropiado para crear un pequeño sistema computador de este tipo. Todos los sistemas computadores pequeños tienen posibilidades, en el campo de soporte lógico, similares en complicación a las que pueden encontrarse en sistemas convencionales mucho mayores. Proporcionan también un acoplamiento para tubo de rayos catódicos o consola con teclado y visualización. Además, gran parte de los sistemas pequeños puede conectarse a dispositivos periféricos tales como memorias en disco flexible, unidades cassette de cinta magnética, unidades de cinta de papel e impresoras de líneas. Con tales mejoras un pequeño sistema computador puede desempeñar las tareas de un sistema completo de desarrollo. En su mayor parte, sin embargo, los ordenadores de este tipo encuentran hoy su principal mercado entre los aficionados a los computadores, que los utilizan para trabajos de programación sencillos, preparación de textos, cálculos generales y juegos de entretenimiento.

En el nivel siguiente de la jerarquía encontramos los sistemas completos de desarrollo, FDS (del inglés "full development system"). Quizá sea su papel más importante el de proporcionar medios rápidos y eficaces para el desarrollo de módulos microprocesadores baratos, que más tarde puedan fabricarse en gran cantidad, capacitados para resolver problemas de producción, telecomunicaciones o negocios. Dicho de otro modo, el sistema completo de desarrollo es un sistema microelectrónico con potencia suficiente para servir de herramienta auxiliar en el desarrollo de un sistema definitivo más pequeño.

Mientras que el sistema completo de desarrollo puede representar una inversión de unos 15.000 dólares, incluyendo tanto el soporte físico como el soporte lógico, el sistema definitivo será un microprocesador o un microcomputador cuyo precio rondará los 500 dólares, si no menos, cuando se fabrique en gran cantidad. Por ejemplo, si se quisiera desarrollar un microprocesador que optimice el rendimiento de un motor de automóvil (mediante el ajuste continuo de la cantidad de carburante, del tiempo de ignición y de la mezcla de carburante), la unidad final podría ser un pequeño módulo de circuito integrado, fijo a cada motor, y con un coste fácilmente inferior a 100 dólares. Podría usarse un sistema completo de desarrollo para construir los programas que finalmente se colocarían en las memorias ROM y PROM del sistema definitivo. Normalmente, un sistema completo de desarrollo incluye un pequeño sistema computador, un mecanismo impulsor de doble disco flexible con un controlador, una impresora de líneas, un terminal de visualización con tubo de rayos catódicos o una consola teleimpresora, un programador de ROM o PROM y tal vez unos cuantos más dispositivos físicos especializados.

El nivel último de tratamiento del microcomputador es el sistema multiprocesador. El microprocesador representa verdaderamente una computación a bajo precio. Su economía es determinante, hasta el punto de que los microcomputadores se están usando no sólo en muchas aplicaciones en las que la potencia de computador era anteriormente demasiado costosa, sino también en aplicaciones en las que varias docenas de módulos microprocesadores especializados pueden ahora conjuntarse para monitorizar y controlar partes de sistemas industriales o comerciales, ya existentes, en los que el control por computador parecía descabellado. Un conjunto de microprocesadores o microcomputadores así reunidos puede organizarse de dos formas funcionalmente distintas.

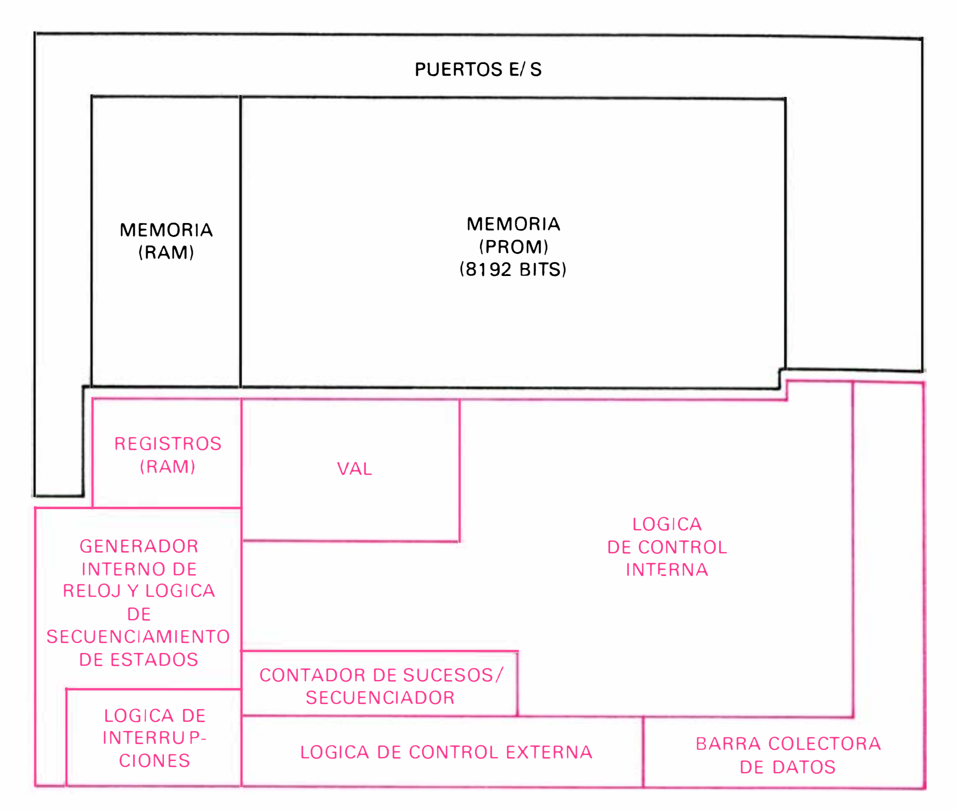
En el primer tipo de organización se diseña un grupo de microprocesadores estrechamente acoplados para intercambiar datos a gran velocidad entre puntos próximos y con alto grado de paralelismo, para alcanzar un máximo de potencia de computación. Un sistema tal podría servir para simular un computador grande, para conseguir alta fiabilidad o para tratar un problema específico que pueda aprovechar la disponibilidad de varios procesadores que trabajen en paralelo.

La segunda organización, que presen-

LOS COMPONENTES BASICOS de un sistema computador pueden incluirse en una sola pastilla, como la 8748 de Intel. En este diagrama de bloques, el "control" abarca la lógica de control y las instrucciones para decodificar y ejecutar el programa almacenado en la "memoria". Los "registros" suministran al control un almacenamiento temporal en forma de memorias RAM (de acceso aleatorio) y funciones asociadas. La unidad aritmética y lógica (ALU) lleva a cabo operaciones aritméticas y lógicas bajo supervisión del control. "Los puertos de I/O" proporcionan acceso a dispositivos periféricos tales como teclado, terminal de visualización, disco flexible ("floppy") para almacenar información e impresora de líneas. Las funciones que, en la ilustración, aparecen en color negro convierten un microprocesador *(color)* en un microcomputador completo.

ta con gran diferencia el mayor potencial de aplicaciones, es un sistema poco acoplado; en él, varios microcomputadores ampliamente distribuidos se comunican datos a baja velocidad con poco o nulo paralelismo. Como ejemplos de tales sistemas distribuidos cabe aducir la automatización de fábricas, control de refinerías de petróleo y plantas químicas y control de dispositivos eléctricos domésticos. Los sistemas distribuidos se están convirtiendo en una realidad a gran escala.

Como se puede imaginar, las herramientas de diseño y desarrollo de sistemas multiprocesadores son mucho más rudimentarias que las de sistemas de procesador único. Los problemas del soporte lógico, cuya resolución no es fácil en un microcomputador, crecen en dificultad, en forma casi geométrica, conforme se van añadiendo unidades al sistema. Estos problemas abarcan la organización de ficheros distribuidos, o sistemas de almacenamiento de información, la planificación de procesos y la realización de lo que a los programadores les agrada llamar "degradación progresiva", lo cual significa que el sistema iría fallando gradualmente en etapas "suaves". Por lo que se refiere al soporte físico del problema, los fabricantes han prestado hasta ahora escasa atención a la configuración de los microcomputadores que les podría haber llevado eficazmente a instalaciones distribuidas.



ESQUEMA DEL MICROCOMPUTADOR 8748 que señala la situación de las diversas funciones del computador. Se repite aquí la disposición de colores utilizada en la figura anterior. Cada función puede ser asignada a uno de los cinco bloques funcionales básicos: control, memoria, registros, ALU y puertos de I/O. Las partes de la pastilla recuadradas en negro representan las funciones que transforman al 8748, de simple microprocesador, en un microcomputador. El dispositivo contiene unos 20.000 transistores fabricados con tecnología MOS *n* (siglas convencionales de la expresión "óxido metálico semiconductor de puerta de silicio de canal *n*"). El procesador central, dotado de 8 bits, maneja 96 instrucciones de 2,5 microsegundos por tiempo medio de ciclo.

Hasta el momento, he utilizado la expresión soporte lógico sin apenas detenerme en su significado. Como quiera que el manejo del soporte lógico es fundamental en todo tratamiento mediante ordenador, voy a ser algo más explícito. Del mismo modo que el soporte físico antes descrito, también el soporte lógico tiene sus jerarquías. En su sentido más amplio, el soporte lógico proporciona los medios para decirle con exactitud a un computador qué debe hacer en cada paso, de acuerdo con la secuencia de instrucciones que constituyen el programa. Cada computador dispone de un "juego de instrucciones": una lista de todas las operaciones básicas que el ordenador puede realizar. Cada una de las instrucciones está escrita en código binario de máquina: una secuencia de ceros y unos, cuya longitud típica es de 8 o 16 bits. Aunque podría escribirse un programa completo en este lenguaje de bajo nivel, resulta un método harto tedioso, y hubo que desarrollar una representación intermedia, conocida por lenguaje ensamblador; éste es el lenguaje más comúnmente empleado hoy en la programación de microprocesadores. Por lo general, cada instrucción simbólica escrita en lenguaje ensamblador representa a una sola instrucción en lenguaje máquina. La traducción la realiza el propio ordenador de acuerdo con un programa de ensamblado.

Para facilitar aún más la programación, se desarrollaron lenguajes "de alto nivel" en los que las instrucciones se aproximan mucho al habla corriente y a las notaciones matemáticas. Ejemplos de tales lenguajes son el FORTRAN, el ALGOL, el COBOL y el PL/1. Una proposición en esos lenguajes suele corresponder a muchas proposiciones en lenguaje máquina. El ordenador realiza la traducción con la ayuda de un programa llamado compilador.

A pesar de tales simplificaciones, resulta complicado escribir un programa. Se llama depurado a la exploración en busca de los inevitables errores y a su corrección. Para permitir la realización de cambios en un programa, el usuario puede utilizar un programa de edición especial, el cual facilita la modificación de instrucciones individuales. Una vez depurado el programa, suele almacenarse en algún dispositivo memorizador estable, como un disco o una cinta magnética. En cuanto esté listo el programa para el ordenador, se transfiere rápidamente desde el disco o la cinta a la memoria de acceso aleatorio y alta velocidad. Como quedó visto, el objetivo normal de un sistema completo de desarrollo es crear un programa que pueda almacenarse en la memoria permanente (ROM o PROM) de un microcomputador final para resolver reiteradamente un problema específico.

A través de los años ha habido una proliferación de lenguajes simbólicos y de alto nivel para propósitos especiales, cada uno de ellos con sus ensambladores o compiladores propios para hacerlos inteligibles en determinados modelos de computadores. A consecuencia de ello, se han desarrollado sistemas de "soporte lógico cruzado" para facilitar la comunicación entre ordenadores. Así pues, los usuarios de grandes ordenadores y de servicios de tiempo compartido tienen acceso a ensambladores de soporte lógico cruzado, a compiladores y a simuladores (programas que permiten a nuestro computador modelizar un duplicado del funcionamiento de otro computador). En este momento, ello representa una al-

ternativa costosa a un sistema completo de desarrollo, con vistas a la creación del soporte lógico para un microprocesador. Además, los simuladores de soporte lógico difícilmente replican las velocidades de entrada/salida y de ejecución en tiempo real del microprocesador final.

Con objeto de imaginar el ritmo de introducción de futuras aplicaciones de los microprocesadores, resulta provechoso comprender el proceso mediante el cual suelen desarrollarse las aplicaciones. Dada una utilización identificable de un microprocesador dentro de un producto o un sistema, ¿de qué forma se desenvuelve una empresa de fabricación para desarrollar un dispositivo adecuado, provisto de un programa adecuado? Como quedó señalado, el usuario potencial no trata normalmente de diseñar una pastilla especial de microprocesador para su tarea específica. Comienza con una de las pastillas existentes ya en el mercado y elige de entre la amplia gama de ellas disponibles: RAM, ROM y PROM, acopladores I/O, o lo que necesite para construir un módulo capaz de llevar a cabo la tarea que desea realizar. En muchos casos, le bastarán los módulos microcomputadores en una sola placa disponible comercialmente y de aplicación general. Los microprocesadores han llegado a abaratarse tanto que suele costar menos dinero aprovechar incluso tan sólo un 5 por ciento o un 10 por ciento de la potencia de computación de una pastilla o un módulo existente, que invertir en el diseño y programación de una unidad especial que realice ese mismo trabajo con el número máximo de componentes electrónicos. Supuesto un volumen suficiente de producción, puede estar justificado, sin embargo, el desarrollo de una unidad especial.

El desarrollo de una aplicación de microprocesador empieza con la identificación de las necesidades. A menudo, la persona de la empresa que entiende de las necesidades no está familiarizada con los detalles de la nueva tecnología microelectrónica. En consecuencia, las necesidades suelen comunicársele a un director de proyectos de ingeniería, el cual evalúa si está justificado o no el empleo de un microprocesador. Tal evaluación podría incluir un análisis para determinar si alguno de los microprocesadores disponibles en el mercado reúne las debidas condiciones y para hacer una estima-

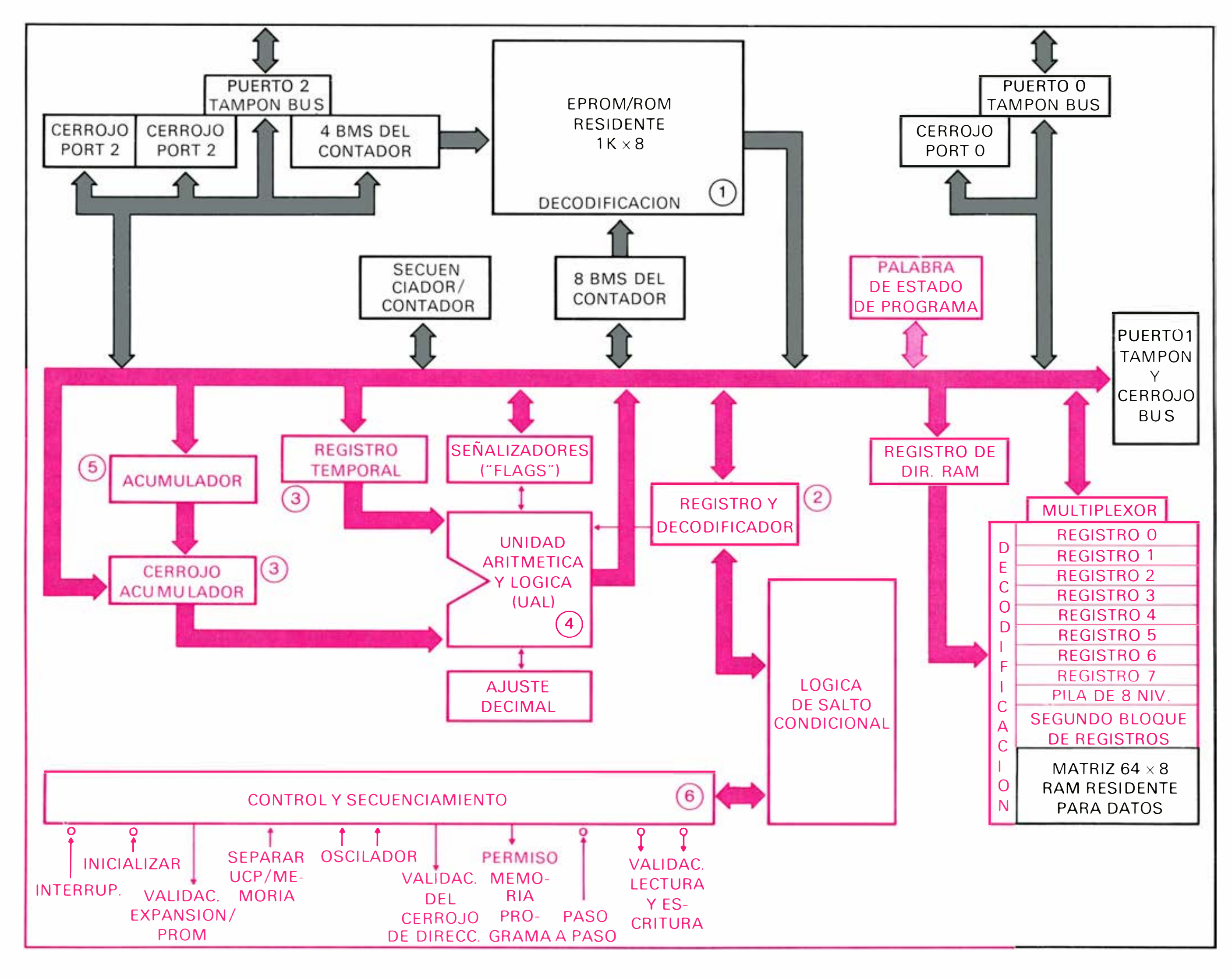


DIAGRAMA FUNCIONAL DE BLOQUES del microcomputador 8748; puede utilizarse para seguir la secuencia de pasos que comporta una operación sencilla, por ejemplo, sumar los contenidos de dos registros *A* y *R*, donde *A* es el acumulador y *R* es uno cualquiera de los registros del bloque inferior a la derecha. El primer paso del computador *(1)* es ir a buscar la instrucción en la memoria: "AÑADIR A, R". El paso siguiente consiste en enviar la instrucción al registro de instrucción y decodificador de instrucciones *(2)*, en donde el decodificador descubre que la instrucción es sumar los contenidos de *R* y de *A* y depositar el resultado en *A*. En el paso siguiente, se envía el contenido del registro *R* al registro temporal *(3)* y el contenido del acumulador al cerrojo del acumulador *(3)*. La ALU *(4)* suma entonces los contenidos de los dos registros y el resultado se reenvía al acumulador *(5)*. Termina la instrucción y se genera una señal *(6)* para ir en búsqueda de la instrucción siguiente.

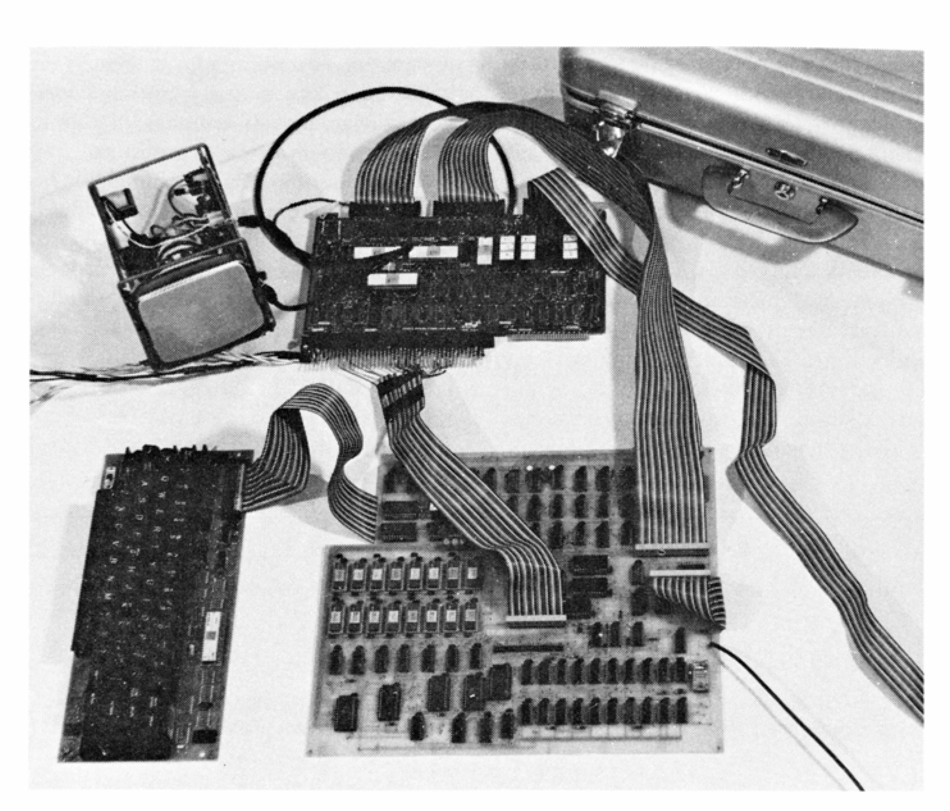
ción del tiempo y del personal precisos para desarrollar el soporte lógico requerido, que es, con mucho, la parte del trabajo que más tiempo y más dinero va a consumir.

Si resulta que está justificada la elección de un microprocesador, se divide el trabajo en dos vías diferentes: los requisitos de soporte físico se traspasan a un ingeniero de diseño y los requisitos de computación y de control se pasan a un programador de soporte lógico. En un caso típico, las tareas comprometidas con el soporte físico y con el soporte lógico se llevan en paralelo mediante dos grupos diferentes. La clave para que resulte una aplicación satisfactoria radica en la estrecha comunicación entre ambos grupos, conforme va gestándose el sistema. Por desgracia, no existe una metodología fija para lograr un buen diseño. Como sucede en toda labor de ingeniería, depende mucho de la intuición, de un buen conocimiento del funcionamiento de los productos disponibles y de la experiencia.

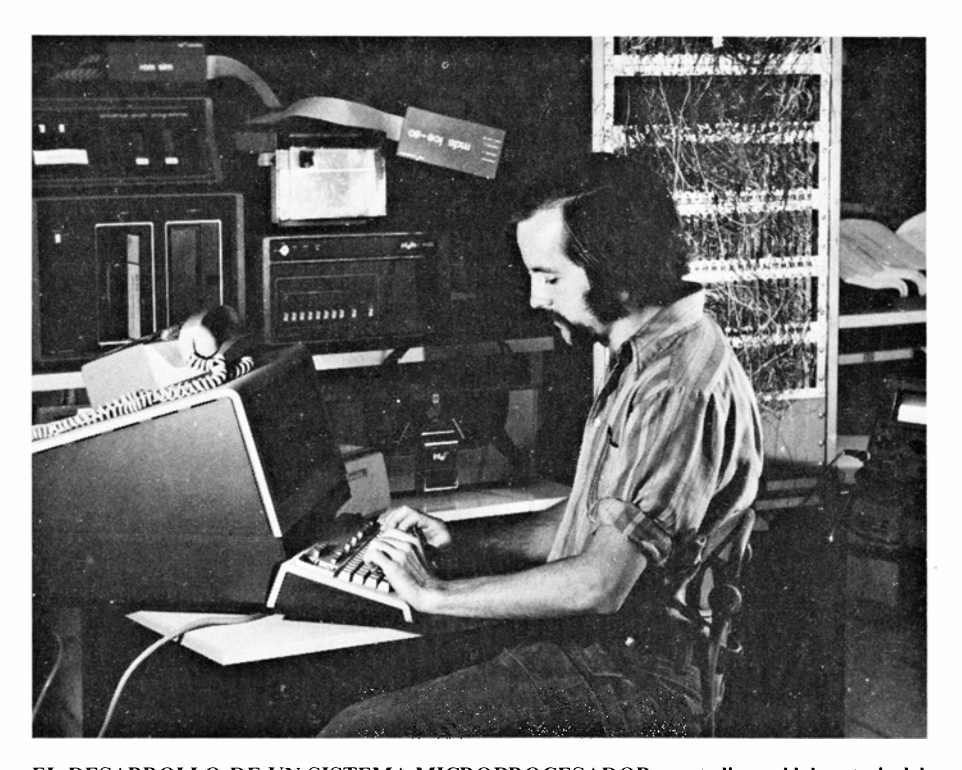
Tanto para el soporte físico como para el soporte lógico, se desarrolla un prototipo del sistema con procedimientos de diseño "ad hoc". Los mecanismos para tener un soporte físico prototipo suelen incluir modelos "en panel" con conexiones por arrollamiento, paneles con tablero de conexiones o prototipos en circuito impreso. El correspondiente prototipo de soporte lógico del sistema consiste en un programa objeto que puede construirse a través de diversos mecanismos: ensamblado manual (mediante una tabla de códigos), uso de un ensamblador residente sobre un sistema completo de desarrollo o empleo de un ensamblador cruzado sobre un sistema de tiempo compartido.

En cuanto se hallan suficientemente avanzados los prototipos de soporte físico y de soporte lógico, hay que reunirlos ambos, tarea que normalmente realiza un ingeniero de sistemas. El programa objeto se carga en el prototipo de soporte físico para comprobar si el sistema resultante cumple las especificaciones originales del director del programa. Las deficiencias a este nivel, en el soporte físico o en el soporte lógico, deben realimentarse a través de "bucles" del soporte físico o del soporte lógico de forma iterativa, hasta que se alcance un funcionamiento satisfactorio.

En este momento debe tomarse una decisión fundamental: poner o no en la cadena de producción al nuevo módulo microprocesador. El período de tiempo transcurrido desde el comienzo del desa-



LOS MODULOS COMPUTADORES en una sola placa proporcionan la oportunidad de diseñar y desarrollar sistemas para tareas especiales. La fotografía muestra un microcomputador "de maletín" construido por el grupo del autor en el Instituto de Tecnología de Massachusetts. Un módulo comercial en una sola placa, un SBC 80/10 de Intel *(junto al maletín)*, se combina con un módulo diseñado por el usuario *(abajo, a la derecha)* y dos unidades periféricas: un teclado y una unidad de visualización. El conjunto demuestra de qué forma sencilla pueden llegar a combinarse módulos normalizados con otros de usuario para la resolución de un problema particular.



EL DESARROLLO DE UN SISTEMA MICROPROCESADOR se estudia en el laboratorio del autor en el Instituto de Tecnología de Massachusetts. El ingeniero de desarrollo está trabajando en un terminal enlazado con un sistema completo de desarrollo del tipo descrito en la ilustración de las dos páginas siguientes. El microcomputador en si, un Intellec MDS-80, está en el armario del fondo, inmediatamente a la izquierda de la cabeza del ingeniero. Justamente debajo del microcomputador hay una lectora de cinta de papel. A la izquierda del microcomputador, se observa un doble disco flexible, y encima hay un programador de PROM. El utillaje de soporte lógico disponible por el ingeniero puede extenderse más allá del sistema completo de desarrollo; es decir, puede incluir habilidades completas del soporte lógico de un minicomputador en sistema "Triad".

rrollo hasta la obtención de un prototipo satisfactorio suele ser crítico. Muchos prototipos se abandonan en esta etapa debido a que el desarrollo ha durado tanto tiempo (puede llevar dos años) que han cambiado los requisitos originales del problema, ha cambiado la idea del problema o un competidor ha podido llegar antes al mercado con un producto equivalente, de modo que se hace necesario ahora salir con algo mejor. Dado ese carácter crítico del tiempo, los grupos de diseño del soporte lógico y soporte físico tienen urgente necesidad de una herramienta de desarrollo que sea más potente que las descritas y que haga posible abreviar, a unas semanas, el tiempo de desarrollo de sistemas prototipo.

Voy a describir una herramienta de esta índole. Como ya hemos visto, el desarrollo de una aplicación del microprocesador prototipo obliga a los diseñadores del soporte físico y a los programadores del soporte lógico a trabajar a casi todos los niveles de las jerarquías que han ido evolucionando en sus disciplinas respectivas. Los diseñadores del soporte físico deben trabajar desde el nivel de pastillas hasta el de sistemas completos de desarrollo; por su parte, los programadores deben trabajar desde el nivel del juego de instrucciones del computador hasta el de sistemas operativos en disco flexible con los diversos compiladores, ensambladores y editores asociados a un sistema completo de desarrollo. Progresar de nivel en nivel dentro de las dos jerarquías es inevitablemente ineficaz, y ello debido a las incompatibilidades habidas entre niveles y a la dificultad de trasladar resultados hacia arriba o hacia abajo, entre distintos niveles.

CARACTERISTICAS Y POSIBILIDADES	USOS Y USUARIOS TIPICOS	JERARQUIA HARDWARE (SOPORTE FISICO) DE MICROS — NIVEL	REPRESENTACION
COMPUTACION DISTRIBUIDA PROCESADO PARALELO FUERTEMENTE ACOPLADO	AUTOMATIZACION PROCESOS COORDINACION Y CONTROL SOBRE UNA BASE DISTRIBUIDA Y LOCAL	SISTEMA MULTIPROCESADOR	SCD PSC O MODULO
DESARROLLO COMPLETO DEL SOFTWARE DEPURADO HARDWARE PROGRAMACION EN LENGUAJE ALTO NIVEL	PROGRAMACION DE APLICACIONES DEPURADO DEL SISTEMA FINAL HARDWARE	SISTEMA COMPLETO DE DESARROLLO (SCD)	DISCO FLEXIBLE DOBLE PROGRAMADOR DE PROM TERMINAL VIDEO LECTORA CINTA PAPEL PEQUEÑO COMPUTADOR
PROGRAMAS DE APLICACION DE COMPLEJIDAD MEDIA (1-10K) POSIBILIDAD ALGUN LENGUAJE ALTO NIVEL (P. EJ. BASIC)	COMPUTADOR PARTICULAR AFICIONADO	PEQUEÑO SISTEMA COMPUTADOR (PSC)	
PEQUEÑO SISTEMA DE DESARROLLO PARA APRENDIZAJE CARACTERISTICAS MICROPROCESADOR PEQUEÑOS PROGRAMAS DE USUARIO (MENOS DE 1 K)	UTILIZADORES PRINCIPIANTES DE MICROPROCESADOR REALIZACION DE PROTOTIPOS ELEMENTALES EVALUACION DEL MICROPROCESADOR POR EL USUARIO	MODULOS	
DISEÑO POR USUARIO DE UN SISTEMA HARDWARE PARA PROBLEMA PARTICULAR	DISEÑADORES HARDWARE	PASTILLAS	

LOS SISTEMAS MICROPROCESADORES pueden disponerse según una jerarquía ascendente de soporte físico y soporte lógico en la que los componentes menores se reúnen en forma de sistemas progresivamente mayores con características más potentes. Los bloques constitutivos son las familias de pastillas diseñadas para funciones diversas. Para resolver un problema concreto de aplicación, por ejemplo el sistema de control de un aeroplano, los diseñadores suelen ensamblar módulos o pequeños sistemas computadores y les dotan de un programa adecuado

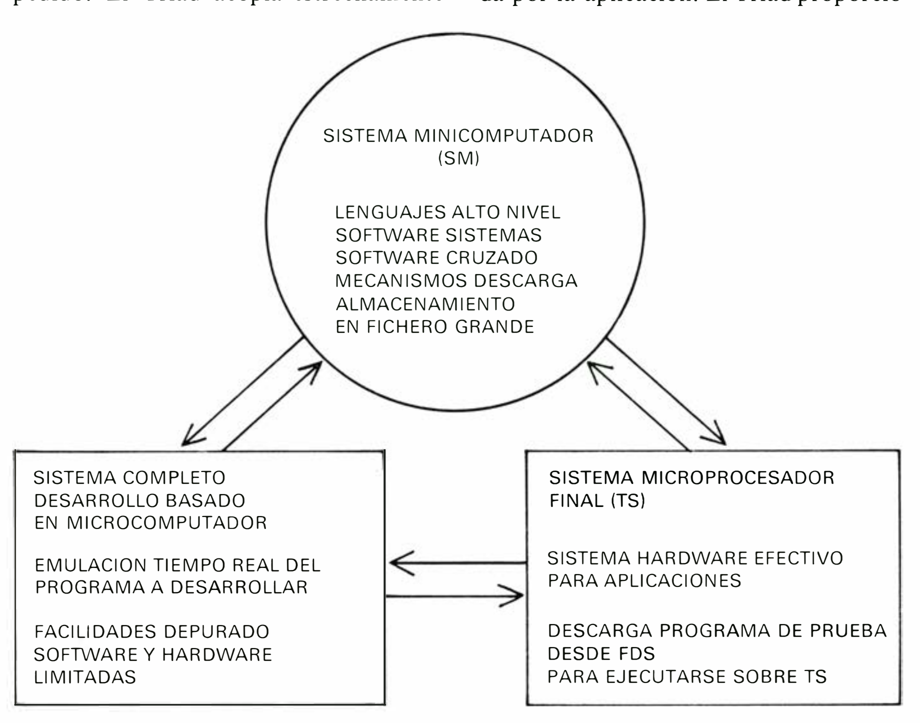
Así pues, ni siquiera los sistemas completos de desarrollo proporcionan todas las características pertinentes que desearían muchos diseñadores de sistemas, verbigracia: amplias facilidades para manejo de ficheros, potentes programas de edición y una biblioteca de soporte lógico para microprocesadores de distintos fabricantes. Los diseñadores que precisan tales exigencias optan por emplear minicomputadores (que son sistemas de computación pequeños, aunque de elevado rendimiento global, que comenzaron a llegar al mercado a comienzos de los años 60), como herramienta complementaria de desarrollo.

Aun cuando existen algunos paquetes de soporte lógico cruzado que permiten desarrollar programas de microprocesador con minicomputadores, hay muy pocos sistemas para acoplar estrechamente minicomputadores, sistemas completos de desarrollo y el sistema final microcomputador. Es deseable un acoplamiento de ese tipo, pues permite al ingeniero de desarrollo moverse con facilidad dentro de esta jerarquía, así como aprovechar al máximo las características diferenciales de cada sistema: el minicomputador, para proporcionar una edición eficaz, almacenamiento masivo, herramientas para documentar y bases de datos compartidas; el sistema completo de desarrollo, para imitar la operación en tiempo real del microcomputador según van construyéndose programas y para depurar el soporte físico; y el propio sistema final microcomputador, para evaluar los programas y rutinas de control finales en las condiciones reales de limitaciones eléctricas y ambientales del montaje en la aplicación.

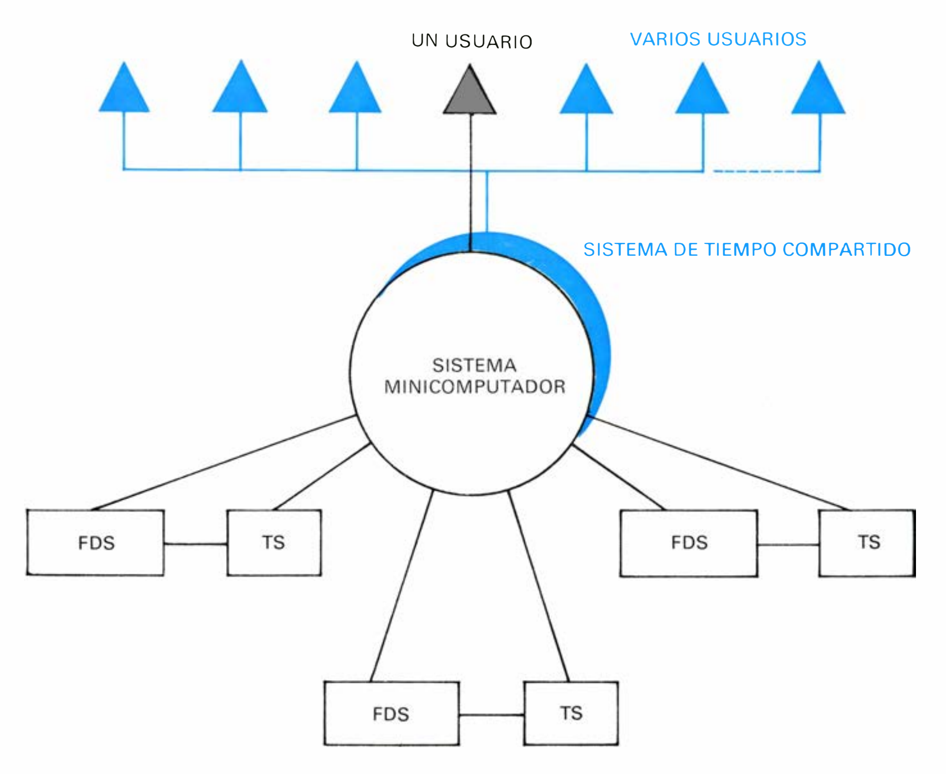
	JERARQUIA SOFTWARE (SOPORTE LOGICO) DE MICROS		
COMPONENTES	NIVEL	REPRESENTACION	COMPONENTES
SISTEMAS MICROCOMPUTADORES SUBSISTEMAS DE COMUNICACION INTERFACES, PARA CONTROL T. REAL, CON SENSORES Y ACTUADORES	SOFTWARE DE SISTEMAS DISTRIBUIDOS	MODULOS RED DISTRIBUIDA PARA MANEJO FICHEROS, CONTROL DISPOSITIVOS, APLICACIONES DE COMUNICACION, ETC.	SISTEMA OPERATIVO DISTRIBUIDO
MICROCOMPUTADOR, TERMINAL VIDEO, DISCO FLEXIBLE, PROGRAMADOR DE PROM, ETC. O SISTEMA SOFTWARE CRUZADO SOBRE MINICOMPUTADOR	SOFTWARE PARA DESARROLLO	SISTEMA OPERATIVO SOBRE DISCO LENGUAJES ALTO NIVEL (P. EJ. FORTRAN, BASIC, PL/M) EJEMPLO: X=X+1 SIENDO X UN NUMERO DE 8 BITS E Y SU DIRECCION	SISTEMA OPERATIVO DISCO FLEXIBLE O CASSETTE COMPILADORES DEPURADORES INTERNOS
VERSION AUTOCONTENIDA DE MODULOS CON POSIBILIDAD EXPANSION MEMORIA E INTERFACES	SOFTWARE AUTONOMO (SIN PERIFERICOS)	LENGUAJES ALTO NIVEL SENCILLOS (P. EJ. BASIC) LENGUAJE ENSAMBLADOR (P. EJ. 8080) LXI H, Y MOV A, M ADI 1 MOV M, A	MONITOR AUTONOMO ENSAMBLADOR EDITOR
TARJETA CIRCUITO CONEXIONES CONTENIENDO FAMILIA DE PASTILLAS Y POSIBILIDAD EXPANSION PARA VARIAS PLACAS	SOFTWARE ELEMENTAL	LENGUAJES SIMBOLICOS ENSAMBLADORES (P. EJ. HEXADECIMAL) 21 Y 7E C6 01 77	MONITOR SENCILLO DEPURADOR LIMITADO
UNIDADES CENTRALES DE PROCESO (CPU). MEMORIAS ACCESO ALEATORIO (RAM). MEMORIAS SOLO LECTURA (ROM). ROM PROGRAMABLE (PROM). INTERFACES ENT.-SAL (E/S) OTRAS PASTILLAS USO ESPECIAL	JUEGO DE INSTRUCCIONES ESCUETO	CODIGO MAQUINA BINARIO 0010 0001 Y 0111 1110 1100 0110 0000 0001 0111 0111	OPERACIONES CABLEADAS O MICROPROGRAMADAS A NIVEL PASTILLA

para la tarea. Las pastillas y los módulos se han abaratado tanto, que un coste normal en la ingeniería de una aplicación es lo que vale el soporte lógico destinado a crear el programa del sistema final. Las mejoras en la tecnología de semiconductores están haciendo resueltamente posible que los sistemas a cada nivel incluyan más posibilidades de las asignadas anteriormente al nivel precedente. Los símbolos y caracteres en color muestran el aspecto de una misma instrucción cuando se escribe en distintos lenguajes de nivel creciente, comenzando con el código máquina binario.

La integración más neta de estos tres sistemas se da en el "Triad", que es una herramienta de investigación recientemente desarrollada; permite al programador o al ingeniero trabajar a cualquier nivel mientras está realizando un programa pedido. El Triad acopla estrechamente tres sistemas: el minicomputador, con sus potentes facilidades de soporte lógico; el sistema completo de desarrollo, con su simulación de tiempo real y sus posibilidades de depurado del soporte físico, y el sistema final, con su estructura definida por la aplicación. El Triad proporciona acceso rápido y directo a todo los niveles de la jerarquía.



LA ESTRUCTURA "TRIAD" reduce el tiempo que necesita el desarrollo de programas de aplicación para un módulo microprocesador o un microcomputador. La utilización del sistema minicomputador centralizado facilita la tarea del programador en su desarrollo del soporte lógico para el sistema final. El microprocesador del sistema final está integrado en la labor desde el comienzo.



LAS ELABORACIONES DEL SISTEMA TRIAD pueden justificarse en grandes sociedades en las que estén desarrollándose cierta cantidad de aplicaciones con microprocesador. El sistema completamente en negro está diseñado para dar a un programador acceso a varios Triad diferentes, dedicado cada uno de ellos a un microprocesador distinto. En el sistema, más potente, de tiempo compartido *(color)*, pueden trabajar simultáneamente ingenieros y programadores en las tareas de desarrollo, en las cuales se encuentren implicados algunos microprocesadores diferentes.

El diseñador del soporte físico y el programador del soporte lógico podrían utilizar el Triad de la forma siguiente: El ingeniero del soporte físico desarrolla el sistema final, que podría constar de un módulo computador en una sola placa, comercial y de uso general, y de un módulo adicional diseñado por él mismo que suministra un acoplador con el sistema grande al que va a destinarse el microprocesador. Para depurar el sistema final necesitará algunos programas de prueba. Tales programas pueden prepararse rápidamente con ayuda del editor en el minicomputador, traducirse al código máquina del microprocesador (con un ensamblador cruzado) y cargarse directamente en el soporte físico final aún no depurado. Conforme va corrigiendo los errores en su soporte físico, el ingeniero irá actualizando continuamente sus programas de prueba. Cada cambio puede volverse a ensamblar y cargar en unos pocos minutos.

Simultáneamente con el desarrollo del soporte físico, el programador puede estar editando, ensamblando y simulando sus programas en el minicomputador y en el sistema completo de desarrollo. Podría mantener todos sus programas en el sistema central de ficheros del minicomputador y compartir una biblioteca de soporte lógico de aplicaciones con otros usuarios del Triad. La integración del sistema puede avanzar gradualmente, ya que las mismas ayudas que ha estado utilizando el ingeniero del soporte físico, para cargar programas de prueba en el sistema final, puede también utilizarlas el programador para descargar el programa final de aplicaciones sobre el soporte físico. Pueden conseguirse cambios en esta fase con facilidad mediante un sencillo proceso de reedición, ensamblado y carga. De este modo, el sistema Triad puede reducir enormemente el tiempo preciso para el desarrollo de una aplicación con microprocesador.

Un único sistema central de minicomputador puede supervisar simultáneamente varios sistemas Triad. Una tal disposición permite al usuario desarrollar paquetes de aplicaciones utilizando varios microprocesadores diferentes y seguir manteniendo todos sus programas, documentación e informes de ingeniería dentro del sistema de ficheros del minicomputador central. Por otra parte, otros usuarios pueden compartir también sus programas haciendo así más eficaz el período de desarrollo. Como paso siguiente, el sistema operativo del minicomputador podría escribirse de nuevo para permitir

que se dé tiempo compartido entre varios usuarios simultáneos. Esto da origen a la variación más potente del concepto de Triad.

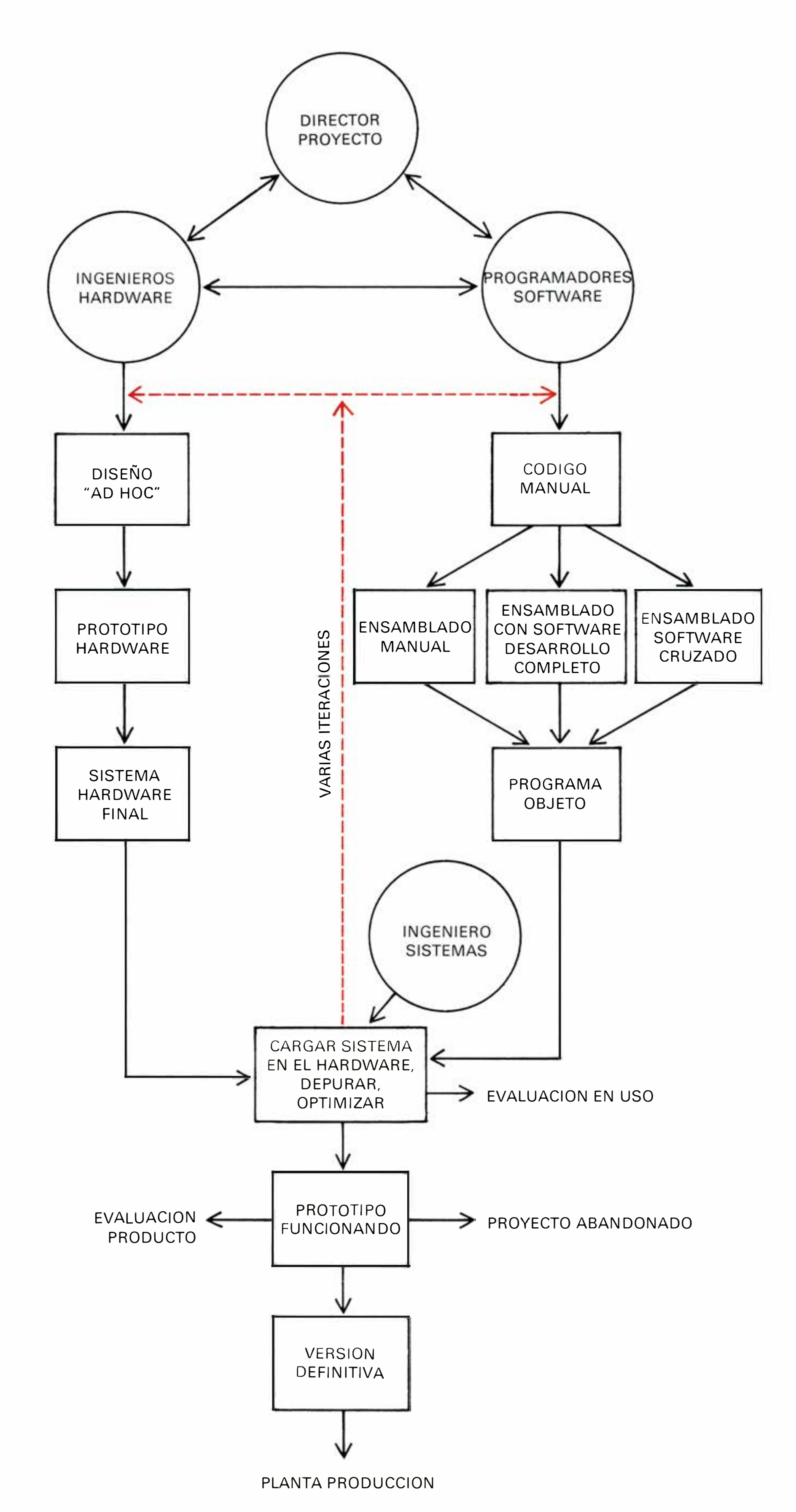
Las aplicaciones potenciales de la tecnología de los microprocesadores son tan numerosas que se hace difícil contemplar cualquier aspecto de la vida actual que pueda escapar a su incidencia. Aquí se presentan algunas de las áreas que serán las primeras en sentir su efecto.

En el sector del automóvil, algunos modelos de 1977 emplean ya microprocesadores para cumplir las leyes que regulan la emisión, controlando los sistemas de recirculación del gas empobrecido. En mayor proporción, los modelos de 1978 incluirán microprocesadores tanto para control de gases emitidos como para optimizar los ajustes del motor con objeto de mejorar el rendimiento de la gasolina. En un futuro próximo, los microprocesadores se conectarán también a los dispositivos de seguridad (por ejemplo, sensores para evitar patinazos sobre superficies húmedas o heladas).

En oficinas comerciales, entre las primeras aplicaciones de los microprocesadores estarán las relativas a la distribución y control de la información. Los "computadores de consola" llegarán a ser casi tan corrientes como las máquinas de escribir. Manejarán bases de datos pequeñas y especializadas, a la medida del trabajo de cada persona, así como información contable y datos de los empleados. La transferencia de documentos mecanografiados entre departamentos se verá en gran parte sustituida por memorándums electrónicos retransmitidos a través del sistema computador de la oficina. En compañías muy importantes, como la International Business Machines Corporation y la Xerox Corporation, están desarrollándose sistemas completos para oficinas.

En la industria, se recurre ya a los microprocesadores para tareas tan diversas como control de máquinas-herramientas y monitorización remota de yacimientos petrolíferos. Los microcomputadores harán también posible una nueva generación de brazos y manos artificiales "inteligentes", capaces de ejecutar operaciones de montaje en fábricas, que hasta el presente resultaban demasiado complejas como para ser mecanizadas.

En el hogar ya han aparecido microprocesadores en una legión de "juegos de video" y aparatos domésticos, como los hornos de microondas y mezcladoras de alimentos. Se extenderán rápidamente a controles térmicos, refrigera-



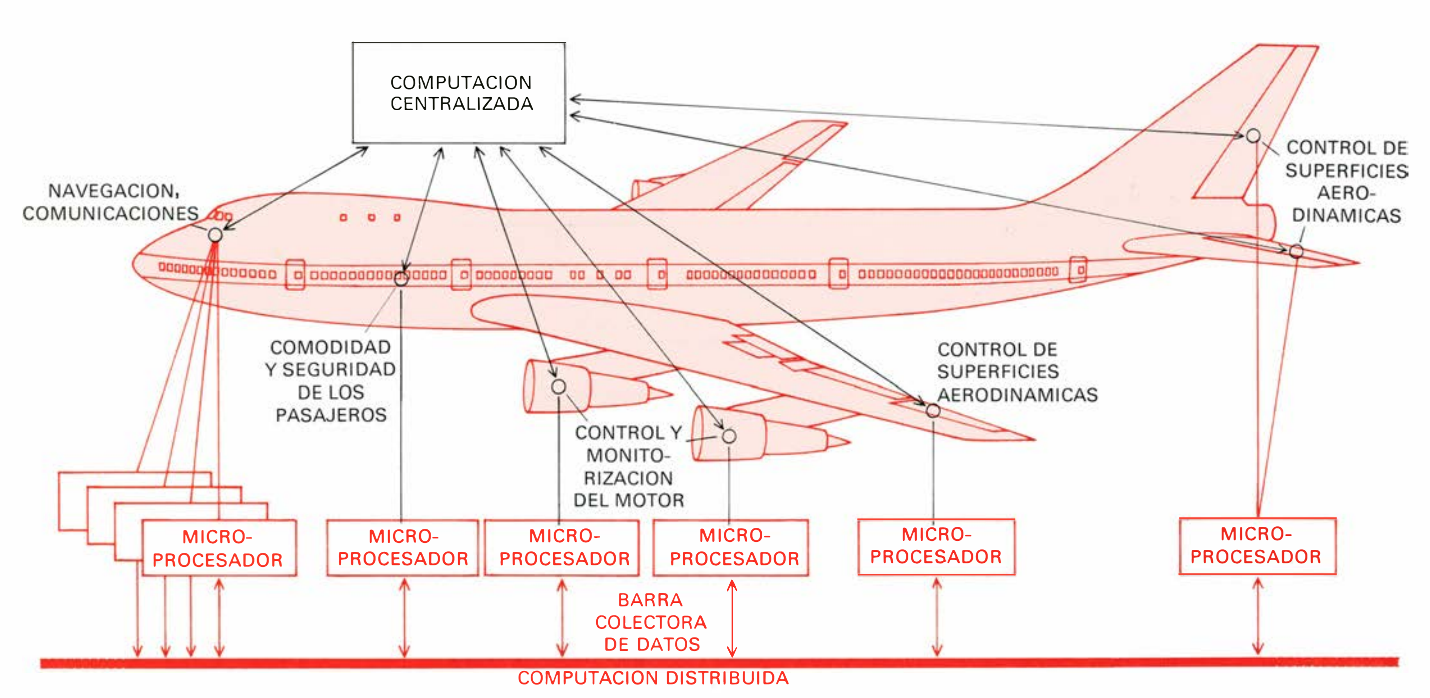
LA LABOR DE DESARROLLO DE APLICACIONES PARA MICROPROCESADORES requiere la cooperación entre ingenieros de soporte físico (hardware), relacionados con la selección y conexión de pastillas, y programadores de soporte lógico (software), cuyo trabajo consiste en proporcionar un programa "depurado" que cumpla los objetivos del proyecto. Como el tiempo para la programación suele estimarse por lo bajo, los objetivos originales han quedado a veces anticuados para cuando está afinado el prototipo y listo para la producción. A consecuencia de lo cual, muchos proyectos se abandonan en esta etapa. No es raro que los programas de desarrollo tarden dos años. El autor cree que puede reducirse notablemente el factor tiempo con el sistema Triad.

dores, teléfonos, sistemas de energía solar y sistemas de alarma contra incendios y robo. Con el tiempo, microcomputadores unidos directamente al banco de cada uno, y dotados con ficheros de varios millones de bits para archivos personales, serán tan comunes como hoy lo son los sistemas sonoros de alta fidelidad.

Idear una aplicación fructífera del microprocesador, sin embargo, es tan sólo el primer paso hacia el logro de su adopción y aceptación. El advenimiento del microprocesador hace presagiar algo más que una mera revolución técnica. Probablemente va a influir en mayor número de aspectos de la vida diaria de los que se han visto afectados por toda la tecnología del ordenador hasta el momento. En muchos casos, la sociedad ni está enterada ni está preparada para el impacto no técnico del microprocesador. Por ejemplo, la introducción de microprocesadores en los automóviles, para aumentar el ahorro de carburante y reducir las emisiones de escapes, tendrán un profundo efecto sobre decenas de miles de pequeños talleres de reparación de automóviles y sobre cientos de miles de estaciones de servicio de gasolina. Millones de obreros dedicados al mantenimiento, que en este momento puede que jamás hayan oído hablar de microprocesadores, y mucho menos los hayan visto, habrán de familiarizarse rápidamente con ellos y convertirse en expertos en su comprobación y reemplazamiento. De otro modo, habría que remozar drásticamente toda la red de servicio de automóviles.

En otras sectores, la tecnología del microprocesador se desenvolverá con menos rapidez de lo que cabría esperar. Por ejemplo, no sería difícil equipar a cada estación de servicio de gasolina con un terminal microprocesador que registrase los detalles de todas las transacciones con tarjetas de crédito. Al final de cada jornada, la información registrada podría transmitirse inmediatamente en código sobre un circuito telefónico normal a los ordenadores de la oficina central de la compañía de tarjetas de crédito, adelantando así en varias semanas las facturaciones de ventas por valor de millones de dólares. Aunque tales sistemas son completamente practicables, no han sido adoptados. Podría deducirse que la innovación no es bien recibida porque amenaza la existencia de departamentos completos de organizaciones de tarjetas de crédito: los departamentos de reconocimiento óptico de caracteres (OCR, del inglés "optical - character - recognition") que transcriben hoy en forma utilizable por el computador los datos de las transacciones a partir de millones de hojas distintas de papel rellenadas y echadas al correo por empleados de estaciones de servicio. No desaparecen de la noche a la mañana, alegremente, secciones enteras de una empresa.

Hemos avanzando dos ejemplos de cómo pueden verse afectados miles de organizaciones comerciales y millones de individuos por lo que parece ser una decisión ingenieril directa sobre si aplicar o no aplicar la nueva tecnología del microprocesador. Dentro de la propia organización comercial, el microprocesador y el microcomputador están agudizando la ya difícil cuestión de cómo distribuir los recursos computacionales e informáticos para lograr un máximo de efectividad. Si bien muchas empresas están consolidando y centralizando con rapidez todos los aspectos de sus recursos de computación, otras empresas los están descentralizando con igual agresividad. La evasiva estrategia "óptima" implica consideraciones delicadas sobre el control de gestión, estrategias de desarrollo de sistemas y procedimientos operacionales. Los directores comerciales que tuvieron que penetrar en los misterios del computador de gran tamaño (room-sized), de 250.000 dólares, hace apenas 15 años y del minicomputador de 25.000 dólares hace seis o siete años, tienen ahora que plantearse la nivelación de costes y beneficios del microcomputador de 250 dólares y el microprocesador de 25 dólares.



LOS MODERNOS AVIONES A REACCION dependen de varios y complejos sistemas para la navegación, comunicaciones, comodidad y seguridad, control del motor y de las superficies aerodinámicas. Los sensores que supervisan los sistemas transmiten sus datos a un computador central, que genera las señales de control precisas para mantener un funcionamiento correcto. Los kilómetros de cables necesarios para tales sistemas centralizados han llegado a representar una parte significativa del coste total de los aviones modernos. En principio, pueden controlarse localmente, por medio de microprocesadores, diferentes sistemas de navegación aérea, con un gran ahorro en el coste de computación y menores costes de mantenimiento. Tales redes de computación se encuentran en estudio activo para una amplia variedad de aplicaciones similares.

Microcomputadores

Las funciones elementales realizadas por los microcomputadores son muy similares en todos ellos. Las diferencias radican en las soluciones adoptadas al integrar en una o varias pastillas de semiconductor los elementos del computador

Antonio Alabau Muñoz y Joan Figueras Pàmies

La aparición de las máquinas digitales para el tratamiento de la información ha estado fundamentalmente motivada por la existencia de dispositivos físicos (transistores, relés), capaces de materializar de forma fácilmente diferenciable dos valores de una determinada magnitud eléctrica (tensión), por la existencia y el desarrollo de una base matemática (álgebra de Boole), capaz de permitir la formalización de relaciones entre variables que pueden tomar dos únicos valores (variables binarias), y por la relativa simplicidad de materializar, mediante aquellos dispositivos, funciones que relacionan variables binarias (funciones lógicas).

El perfeccionamiento de las técnicas de fabricación de componentes digitales, como se ha podido observar a lo largo de varios artículos de este número, ha dado lugar a la aparición de un importante elenco de unidades funcionales integradas (circuitos lógicos), tradicionalmente clasificadas en tres grandes grupos: unidades con baja densidad de integración –denominados abreviadamente SSI– (puertas lógicas, biestables), unidades con densidad media de integración –MSI– (registros, contadores) y unidades de elevada densidad de integración –LSI– (memorias, funciones especiales, microprocesadores). Cada uno de estos circuitos lógicos estará materializado en una pastilla de semiconductor, caracterizado por la función (o las funciones) que pueda realizar.

Un circuito lógico integrado es un elemento utilizable en la realización de máquinas digitales. No obstante, dado que a su vez es capaz de realizar un determinado tratamiento de la información, aunque en muchos casos sea muy sencillo (operación lógica, contaje, memorización), podremos considerarlo asímismo como una máquina digital.

Toda máquina digital, construida a base de circuitos lógicos, resulta de la agregación de máquinas digitales más elementales; el conjunto obtenido será capaz de realizar una función de orden superior a la de cada una de las máquinas que la componen.

Ante la rápida evolución de la microelectrónica, la revolución de las máquinas digitales ha ido "a remolque" de los avances en la fabricación de los circuitos lógicos, muchas máquinas han quedado rápidamente obsoletas al aparecer nuevos tipos de circuitos integrados.

Los ordenadores y demás máquinas digitales de la década de los sesenta estaban realizadas con componentes discretos (diodos, transistores). Era la época del diseño lógico que utilizaba únicamente un solo tipo de máquina elemental básica: la puerta "no-y" (NAND) o "no-o" (NOR); recordemos que es posible realizar cualquier máquina digital con circuitos lógicos NAND (o NOR) por tratarse de módulos funcionalmente completos, capaces de ejecutar las funciones lógicas "y", "o" y "complemento", suficientes para expresar cualquier función lógica por compleja que ésta sea.

Con la aparición de los primeros circuitos integrados se pasó a construir máquinas digitales a base de módulos funcionales especializados, fundamentalmente: puertas lógicas, sumadores, biestables, registros, multiplexores y demultiplexores.

Finalmente, en la década de los setenta aparece la posibilidad de integrar en un módulo LSI, circuitos complejos, lo que únicamente resulta rentable en los casos de volúmenes elevados de producción de dicho circuito (calculadoras de sobremesa); consecuentemente aparece la necesidad de adoptar una nueva alternativa en la realización de máquinas digitales como solución al problema de compatibilizar la utilización del alto grado de integración en la fabricación de circuitos con la no limitación de las posibilidades funcionales del producto resultante.

La posibilidad de disponer de "una máquina cuya estructura fuera independiente de las funciones que fuera capaz de realizar" constituía una solución muy atractiva. Si, además, dicha máquina podía realizar cualquier tipo de función, sin introducir modificaciones en su estructura, como en el caso de las máquinas de utilización general o computadores, se convertía en la solución ideal. En la búsqueda de aproximaciones a dicha solución ideal, aparecieron los microcomputadores.

Un microcomputador es, pues, una máquina de utilización general realizada en uno o varios módulos de circuito integrado, esto es, una máquina formada por uno o varios módulos LSI, cuya función puede ser programada por el diseñador. En cualquier caso, el microcomputador constituye la solución adoptada para aprovechar las posibilidades de la fabricación de circuitos con una elevada densidad de integración, en la realización de máquinas digitales.

La realización de una máquina digital basada en la utilización de un microcomputador consistirá en un proceso de descomposición de la función que se desea que se realice en un conjunto ordenado de funciones más elementales, capaces de ser ejecutadas por el microcomputador, es decir, en expresar la función a realizar, mediante un algoritmo. La codificación de dicho algoritmo mediante las instrucciones del microcomputador constituirá el programa que se almacenará en su memoria.

Modelos para el estudio de máquinas digitales

Un determinado problema puede ser resuelto bien mediante el diseño de una máquina digital especializada, bien mediante la utilización de un microcompu-

tador. Desde el punto de vista funcional y con independencia de su estructura, ambas soluciones responden a unas mismas características básicas: son capaces de tomar información del exterior, procesarla a base de la ejecución de funciones más elementales y devolver el resultado de dicho tratamiento al exterior.

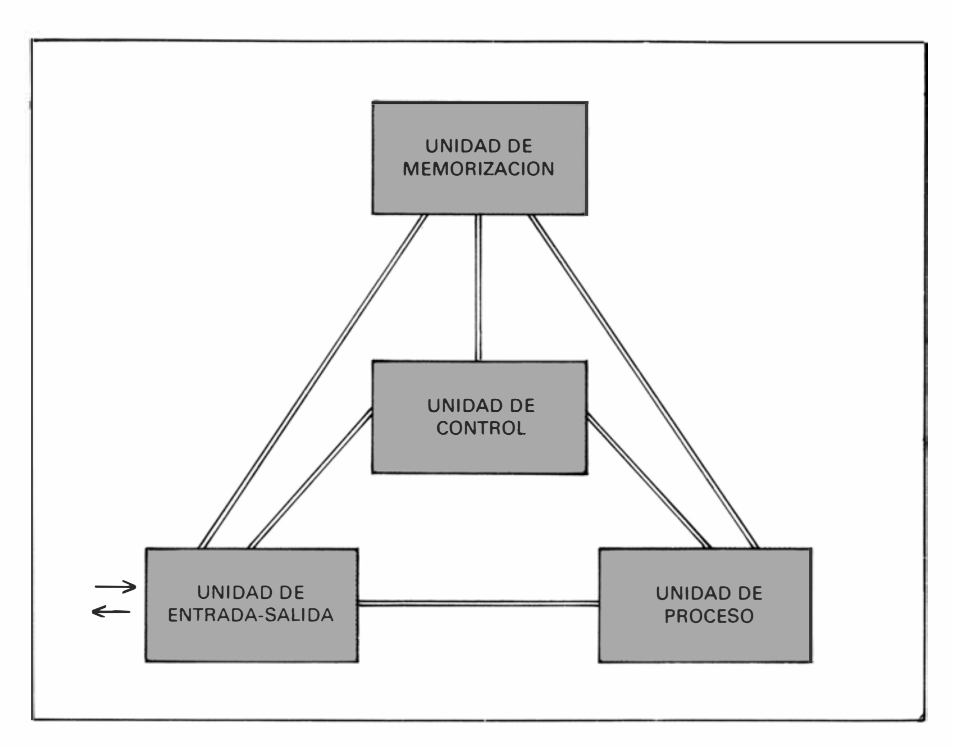
En un intento por estudiar de una forma unificada las características de los diferentes tipos de máquinas digitales, ya sean especializadas, ya sean de utilización general, presentaremos a continuación dos modelos que reflejan aspectos complementarios de las mismas; se trata de un modelo de unidades operativas y de un modelo funcional.

El *modelo de unidades operativas* representa a una máquina digital compuesta de las siguientes unidades funcionales o subsistemas básicos: (1) *Unidad de entrada y salida* de las informaciones; a través de la cual se realizará la comunicación con el exterior para el intercambio de datos y resultados así como para la recepción de órdenes de ejecución de funciones. (2) *Unidad de proceso* o de tratamiento de las informaciones, bien procedentes del exterior bien resultantes de operaciones anteriores. (3) *Unidad de memorización*, que almacenará, tanto los resultados intermedios de las operaciones que ejecutan, como el programa que deberá ejecutarse, en las máquinas de programa almacenado. (4) *Unidad de control*, que ejercerá la coordinación de la correcta ejecución y secuenciamiento de todas las operaciones que se lleven a cabo en el resto de los órganos de la máquina.

La forma en que dichas unidades se materializan en una determinada máquina dará lugar a la estructura de la misma.

El *modelo funcional* representa a una máquina digital mediante un triple conjunto de elementos designado por *D, P, A*; siendo: *D*, el conjunto de *estructuras de información* (y de entrada-salida, interna) que pueden utilizarse en la máquina; *P*, el conjunto de *primitivas* que la máquina es capaz de ejecutar, esto es, el conjunto de órdenes codificadas que pueden darse desde el exterior y que la máquina puede interpretar y realizar (estas primitivas manipularán informaciones contenidas en el conjunto *D*); *A*, el conjunto de *algoritmos* que pueden ser descritos empleando únicamente las primitivas del conjunto *P*, y que representan las posibilidades funcionales de dicha máquina.

Mediante dicho modelo es posible poner de manifiesto las posibilidades funcionales de la máquina desde el punto de vista del usuario.



UN MICROCOMPUTADOR puede estudiarse a través de un modelo de unidades operativas, en el que el conjunto de sus elementos está representado por cuatro unidades o subsistemas básicos: unidades de entrada-salida, unidad de memorización, unidad de proceso y unidad de control. La forma en que dichas unidades se materialicen dará lugar a la estructura del microcomputador.

A lo largo del texto, utilizaremos ambos modelos para analizar las características estructurales y funcionales de distintos microcomputadores.

Características estructurales de los microcomputadores

Los microcomputadores constituyen, pues, una clase de máquinas de utilización general realizadas con circuitos de alta densidad de integración (LSI). Cada microcomputador dispone de una familia más o menos extensa de elementos (memorias, microprocesador, entrada y salida), permitiendo de este modo adaptar y ajustar sus características a las exigencias de las aplicaciones en las que serán utilizados.

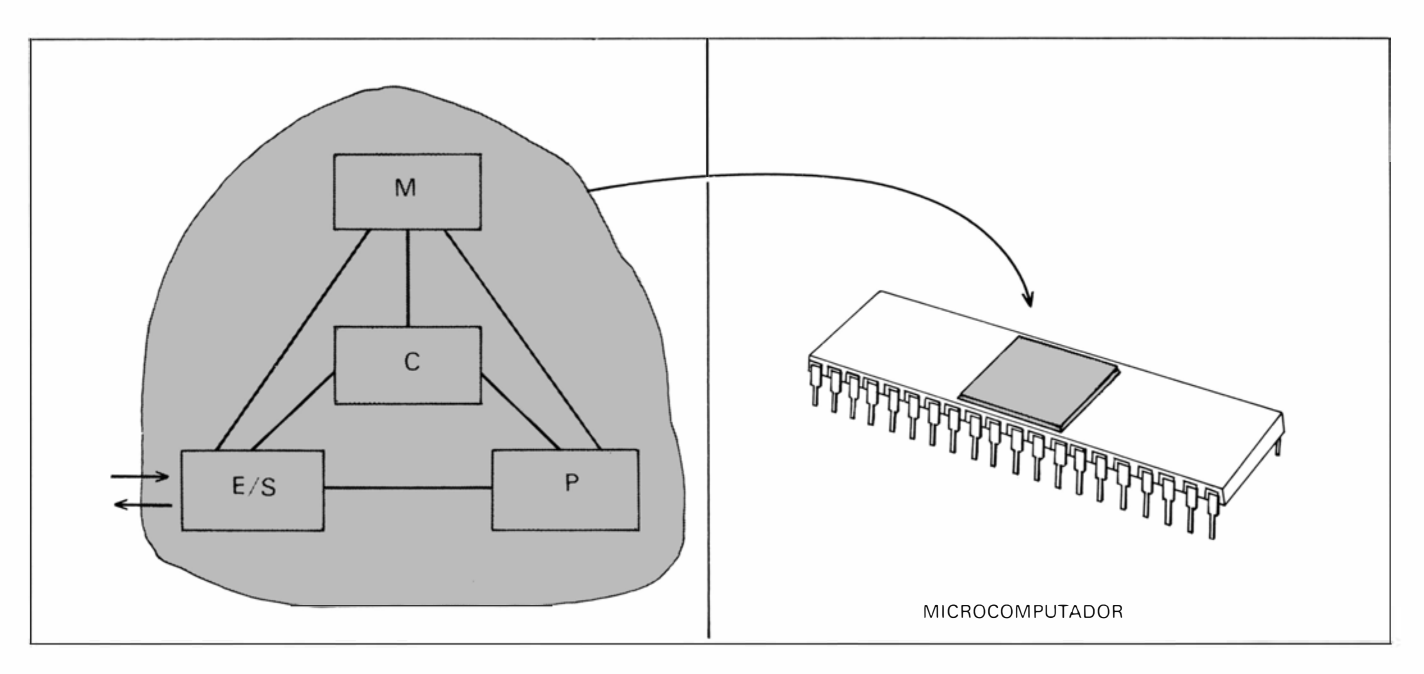
Cada fabricante adopta la solución que le parece más idónea, tanto en lo que se refiere a la estructura de la máquina, como a las posibilidades funcionales de la misma, integrando en bloques independientes aquellas partes que, según su criterio, resulta posible y rentable presentar como única pastilla de circuito.

El fabricante tiene ante sí una doble alternativa: realizar su microcomputador integrando la totalidad de sus unidades funcionales en una única pastilla de circuito integrado o bien hacer una determinada partición en dicho modelo integrando por separado los subsistemas resultantes.

En los párrafos que siguen trataremos de presentar someramente distintas alternativas adoptadas en la realización de un microcomputador; mientras que en unos casos aparece integrado en un circuito la totalidad del microcomputador, en otros, sólo lo está una parte de las unidades de control y proceso, las cuales junto con bloques integrados de memoria y unidades de entrada y salida habrán de constituir la máquina microcomputador.

En nuestra presentación, seguiremos un orden correspondiente a la partición elegida que en ningún modo corresponde con el orden cronológico de aparición de los componentes en el mercado, agrupando los microcomputadores en 3 grandes grupos: microcomputadores integrados en una pastilla, microcomputadores basados en microprocesador y microcomputadores a "rebanadas".

En primer lugar nos referiremos a los microcomputadores integrados en una pastilla. Existen algunos sistemas en los que la totalidad de las unidades operativas que constituyen el microcomputador han sido realizadas en una única pastilla de circuito integrado. Comoquiera que las unidades de control, proceso, memorización y entrada-salida se encuentran presentes en el circuito, éste constituye un conjunto funcionalmente completo, por lo que puede ser utilizado directamente en una aplicación.



EL MICROCOMPUTADOR puede realizarse en una única pastilla de circuito integrado, la cual contendrá la totalidad de las unidades operativas que constituyen la máquina. Se trata de sistemas autosuficientes; por tanto, pueden utilizarse directamente en una aplicación sin elementos complementarios. El TMS 1000, y el TMS 9940 de Texas Instruments, el 8048 de Intel y el 3870 de Mostek responden a estas características.

Las posibilidades de estas máquinas vendrán fundamentalmente determinadas por las características de sus elementos de memorización y de entrada y salida. En la mayoría de los sistemas actualmente existentes, la capacidad de memoria para almacenamiento de programa oscila entre 1024 palabras (1K), y 2048 palabras (2K) de memoria ROM, disponiendo, como máximo, de algunos cientos de palabras de memoria RAM para el almacenamiento de valores variables, lo que limita el tipo de aplicaciones en las que pueden ser utilizadas.

Otro factor que condiciona la utilización de este tipo de componente es la forma en que debe introducirse el programa en la memoria ROM; aquellos sistemas que exijan una programación "por máscara", esto es, una programación en el momento de la fabricación del componente, verán sus posibilidades limitadas a aplicaciones en las que las series de producción sean elevadas y esté justificada la inversión necesaria para la realización de las máscaras de grabación del programa. Por este motivo, uno de los esfuerzos de los fabricantes ha consistido en dotarles de la posibilidad de ser programados (y reprogramados si es necesario) por el usuario, utilizando memorias programables eléctricamente y borrables por radiación ultravioleta, integradas sobre el mismo substrato que el resto de los elementos de la máquina.

Dado el atractivo de este tipo de microcomputadores, es de prever una rápida evolución, tanto en sus posibilidades de memorización como en sus restantes características, fundamentalmente en sistemas compatibles desde el punto de vista del soporte lógico con máquinas existentes en el mercado.

En segundo lugar hablaremos de los microcomputadores basados en microprocesador. Las limitaciones encontradas en los sistemas anteriores, respecto a la reducida capacidad de memorización contenida en la pastilla microcomputador, no aparecen en los sistemas en los que se ha adoptado la solución de integrar las unidades de control y proceso en un circuito independiente de las del resto del sistema (memoria y entrada-salida).

En este tipo de sistemas, se denominará *microprocesador* al circuito que contenga la unidad de proceso del sistema complementado con la parte fundamental de la unidad de control y, en algunos casos, con otro tipo de elementos de los que constituyen el microcomputador. El microprocesador es el componente resultante de la decisión del fabricante de integrar en un solo bloque un conjunto de elementos, sin duda los fundamentales, de los que constituyen un microcomputador.

La comprensión de este tipo de sistemas implica conocer la solución estructural adoptada en la fabricación del microprocesador, saber cómo están constituidos el resto de los elementos que forman el sistema (la memoria y los elementos de entrada/salida) y estudiar la forma de interconexión y los mecanismos que permiten el funcionamiento conjunto de estos elementos que constituyen el microcomputador.

Por lo que se refiere al microprocesador, las diferencias estructurales entre los distintos tipos se concretan en las partes de las unidades de control, proceso, memorización y entrada/salida contenidas en dicho componente.

Según las funciones integradas en el circuito microprocesador, es posible encontrar las siguientes soluciones: una solución en la que el microprocesador no contiene la totalidad de los elementos que constituyen la unidad de control, en el que ciertos aspectos de la generación de señales de control deben realizarse en el exterior del circuito; otra solución en la que el microprocesador contiene la totalidad de los elementos de control, aunque sin ningún elemento de entrada-salida; y una tercera solución en la que el microprocesador, además de las unidades de proceso y control, incluye algunos elementos de entrada y salida.

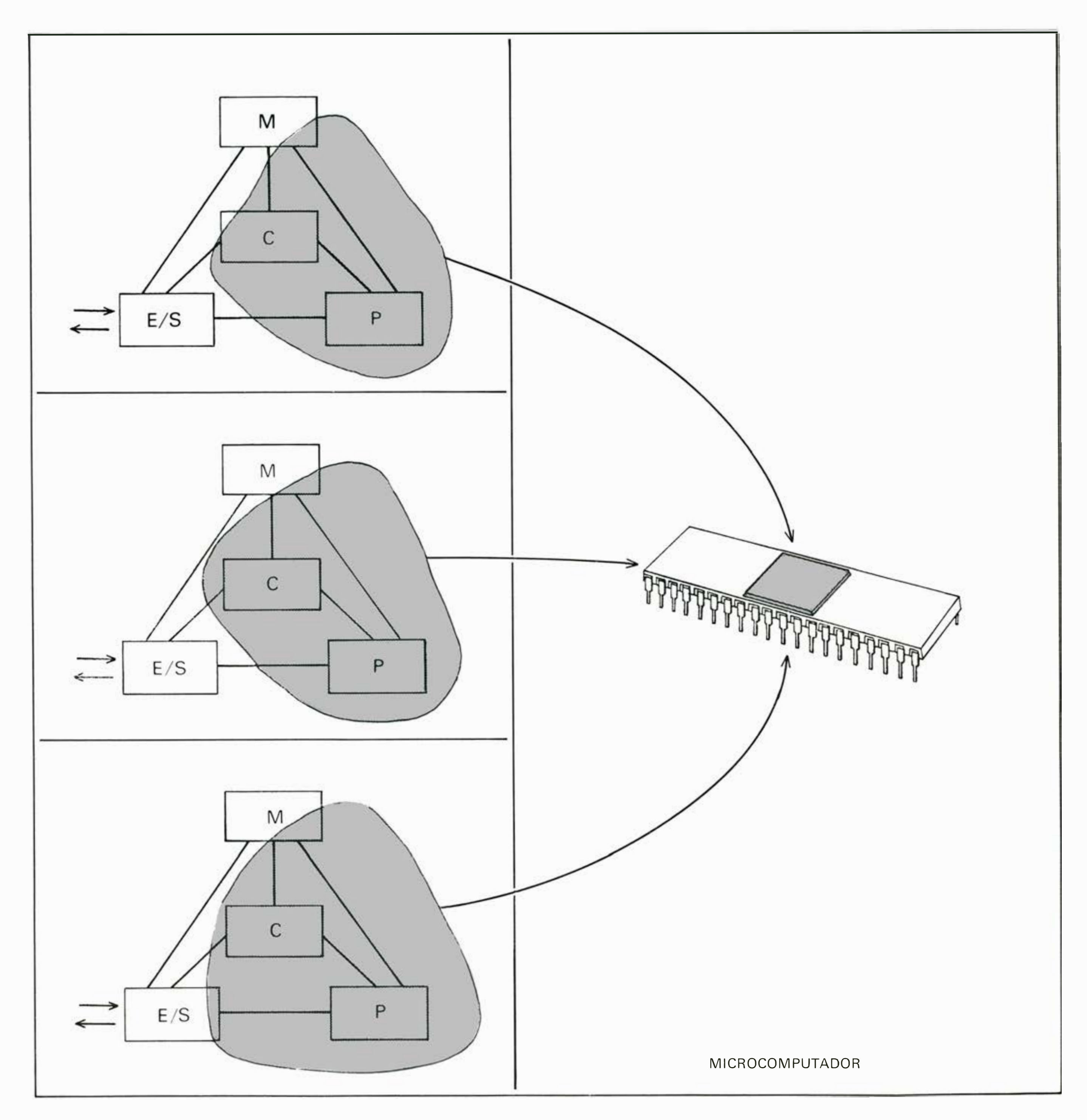
En la ilustración de la página 96 se representan esquemáticamente los dos grandes tipos de soluciones que suelen adoptarse en la integración de la memoria RAM y ROM y en los elementos de entrada y salida.

Una alternativa: se integran por separado funciones distintas, dando lugar a la aparición de bloques de memoria RAM, bloques de memoria ROM y bloques que realizan funciones de comunicación con el exterior, lo que constituye

el caso más frecuente. Este tipo de organización permite una evolución y un dimensionado independiente de cada uno de los bloques funcionales.

Otra alternativa, quizá no tan corriente como la anterior, consiste en integrar conjuntamente funciones de memorización y de entrada-salida dando lugar a componentes RAM/entrada-salida y ROM/entrada-salida. La utilización de este tipo de circuitos está casi exclusivamente justificada en el caso de sistemas reducidos, en los que quizá con uno o dos bloques de circuitos integrados se solucionen simultáneamente las necesidades de memorización y las de comunicación con el exterior.

Queda por explicar el tercer gran grupo de microcomputadores, a saber: los microcomputadores a rebanadas. La necesidad de disponer de microcomputadores de alta velocidad planteó la necesidad de utilizar tecnología bipolar en su fabricación, en oposición a la utilización casi exclusiva de tecnología MOS empleada en los tipos de microcomputadores estudiados. La densidad de integración de transistores bipolares es actualmente inferior a la que es posible obtener integrando transistores MOS. Así pues,



UN MICROPROCESADOR es un circuito que contiene la unidad de proceso, la parte fundamental de la unidad de control, y, en algunos casos, elementos de entrada y salida de un microcomputador. Según las funciones integradas en el microprocesador es posible encontrar, entre otras, las siguientes soluciones: una solución en la que el microprocesador no contiene la totalidad de los elementos de control (por ejemplo, el 8080); una segunda solución en la que contiene la totalidad de los elementos de control, aunque sin ningún elemento de entrada-salida (el 6800, el PPS-8, el 6500 y el Z-80 son ejemplos pertinentes), y una tercera solución en la que el microprocesador incluye algunos elementos de entrada-salida (entre otros podemos citar los modelos siguientes: 8085, el 2650, el SC/MP, el F-8, el PACE, el COSMAC y el 9900).

la limitación existente en el número de componentes activos que es posible integrar en una pastilla impone una limitación de las funciones que dichos circuitos integrados pueden realizar. La solución adoptada por los fabricantes consiste en integrar, en bloques independientes, parte de las unidades de control y proceso, de forma que una vez agregadas pudieran constituir un computador. Las características de estos bloques están limitadas en dos direcciones: en primer lugar en cuanto a la función que realizan y, en segundo lugar, en cuanto a la longitud de las palabras (número de bits) con que son capaces de operar.

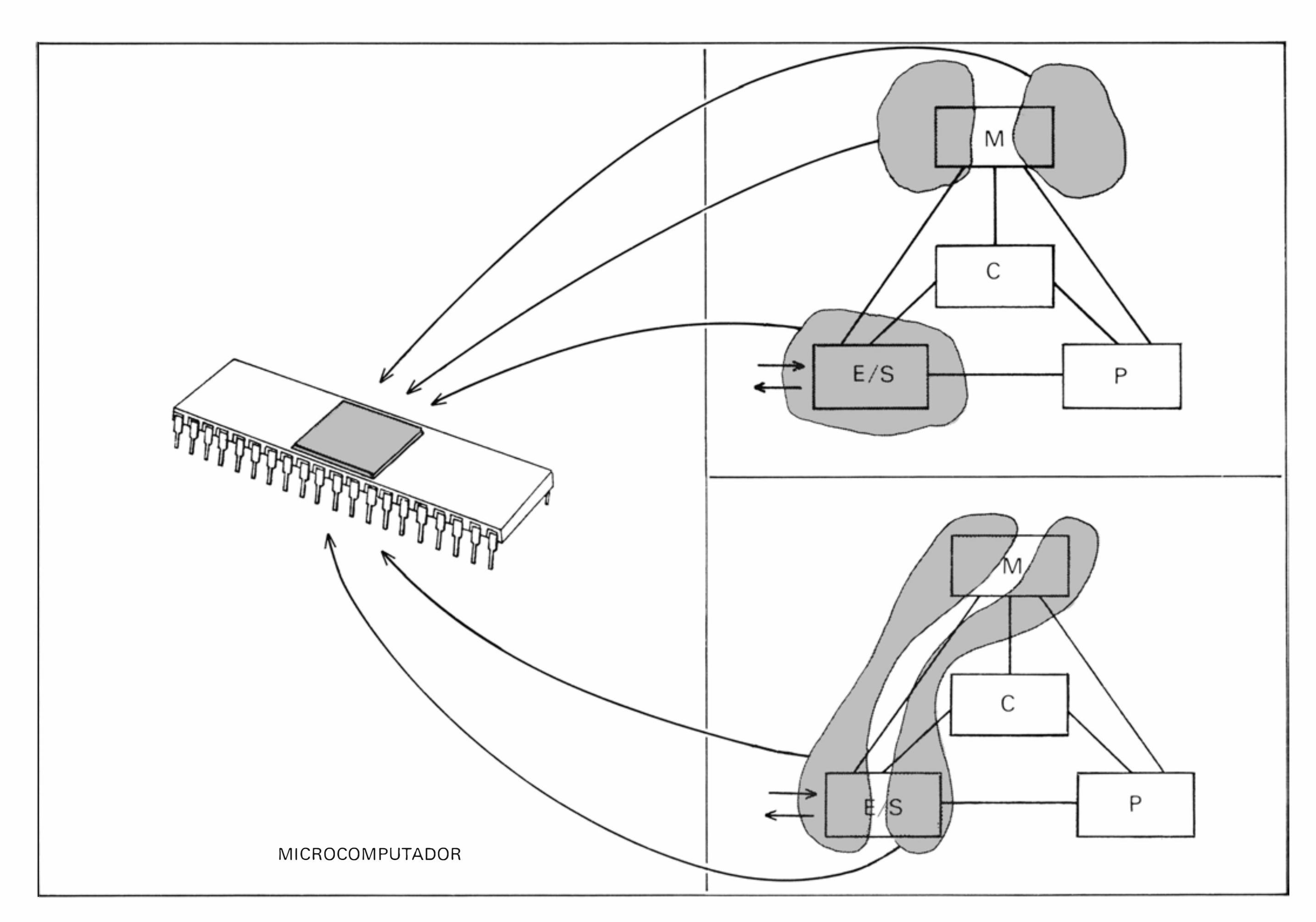
Dado que la unidad de control no constituye un bloque inaccesible para el usuario, sino que está formado por un conjunto de unidades funcionales, el fabricante deja en manos del diseñador la posibilidad de actuar sobre los elementos de interpretación de las instrucciones, y fijar el significado de las mismas, es decir, le da la posibilidad de microprogramar la máquina.

El estudio progresivo de las posibilidades de una máquina justifica la utilización de un modelo que haga abstracción de aquellos detalles innecesarios para la comprensión de sus funciones fundamentales. Aquí emplearemos el modelo funcional presentado anteriormente, según el cual una máquina digital está caracterizada: por las estructuras de información -*D*- de que dispone, por el conjunto de primitivas -*P*- que es capaz de ejecutar, y por el conjunto de algoritmos, *A*, que es posible describir con las primitivas del conjunto *P*, y que representan el conjunto de funciones realizables por la máquina. Lo cual constituye un triple *(D, P, A)*.

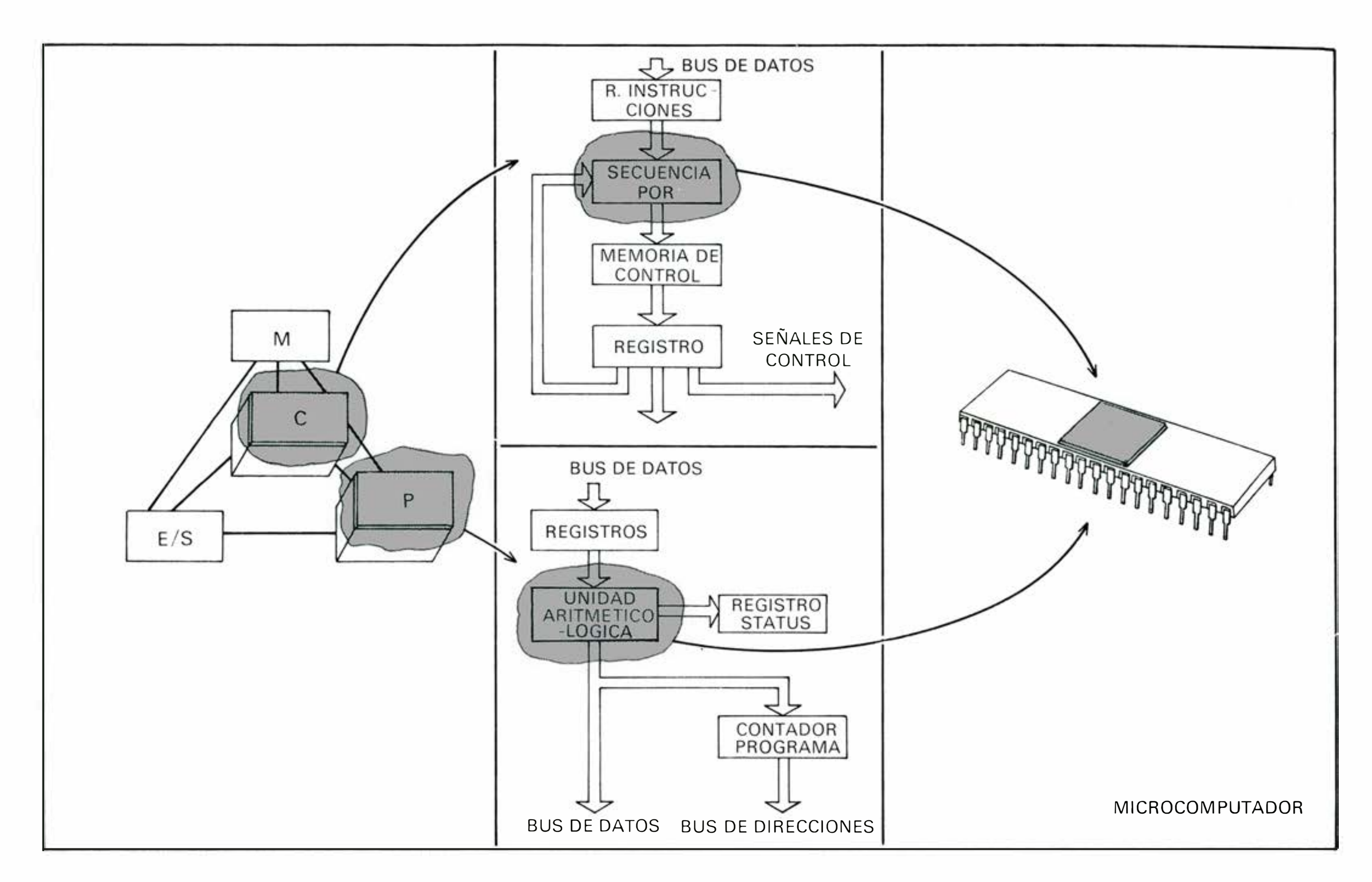
Características funcionales de los microcomputadores

En un microcompuntador, las estructuras de información *D* estarán constituidas por la totalidad de las palabras de memoria existente, los registros del procesador, los registros de entrada y salida, los registros de la unidad de control accesibles al usuario y por todas aquellas estructuras de datos que puedan manejarse utilizando los mecanismos de direccionado del sistema.

El conjunto de primitivas *P* estará constituido, en primer lugar, por los códigos de las instrucciones que el fabricante (o el usuario en algunos casos) ha prefijado y que corresponde a operaciones de tratamiento de la información, de comunicación con el exterior o de control; en segundo lugar, por un conjunto de códigos asociados a funciones que la máquina es capaz de ejecutar, y que indican otras posibilidades del microcomputador además de la ejecución de programas. Están incluidas en este conjunto: las órdenes de iniciación de las operaciones en la máquina (Reset); las órdenes de petición de ejecución de programas situados en determinados lugares de memoria, previa *interrupción* de las funciones que la máquina estuviera realizando; las órdenes de suspensión de las actividades de la máquina, inhibiendo la totalidad, o una parte de ella, del resto del sistema, lo



LOS ELEMENTOS de entrada-salida y de memorización pueden integrarse según distintas alternativas. En unos casos, se integran separadamente bloques de memoria de lectura y escritura (RAM), bloques de memoria de sólo lectura (ROM) y bloques que permiten realizar funciones de comunicación con el exterior; esta solución posibilita una evolución y dimensionado independiente de cada uno de los bloques citados. En otros casos se integran conjuntamente funciones distintas, dando lugar a bloques RAM/entrada-salida y ROM/entrada-salida. Esta solución se utiliza en sistemas reducidos, puesto que permite minimizar el número de componentes necesarios para constituir el microcomputador.



MICROCOMPUTADORES A REBANADAS. Se integran separadamente partes de la unidad de control y partes de la unidad de proceso, las cuales realizan el tratamiento de bloques de bits (rebanadas). Combinando varias unidades de este tipo, se puede aumentar la longitud de las microinstrucciones y la longitud de las palabras con las que el microcomputador opera (el sistema 2900 y el 3000 responden a esta filosofía).

que da la posibilidad de compartir la utilización de determinados elementos del mismo, como en el caso del *acceso directo a la memoria* (DMA), sin la intervención del procesador. Pasarán también a formar parte de esta categoría todas aquellas órdenes orientadas a la utilización conjunta de varios microcomputadores.

Según lo que acabamos de decir, el conjunto de algoritmos *A* que un microcomputador es capaz de realizar dependerá de las posibilidades de combinación de la totalidad de las primitivas mencionadas.

De los elementos del triple *(D, P, A)*, tanto el conjunto de primitivas como las estructuras de información nos muestran un aspecto estático del microcomputador, resumiendo el conjunto de posibilidades del mismo. Por el contrario, el conjunto de algoritmos *A* nos presenta un aspecto dinámico de la máquina, esto es, sus diferentes formas de comportamiento para la resolución de problemas.

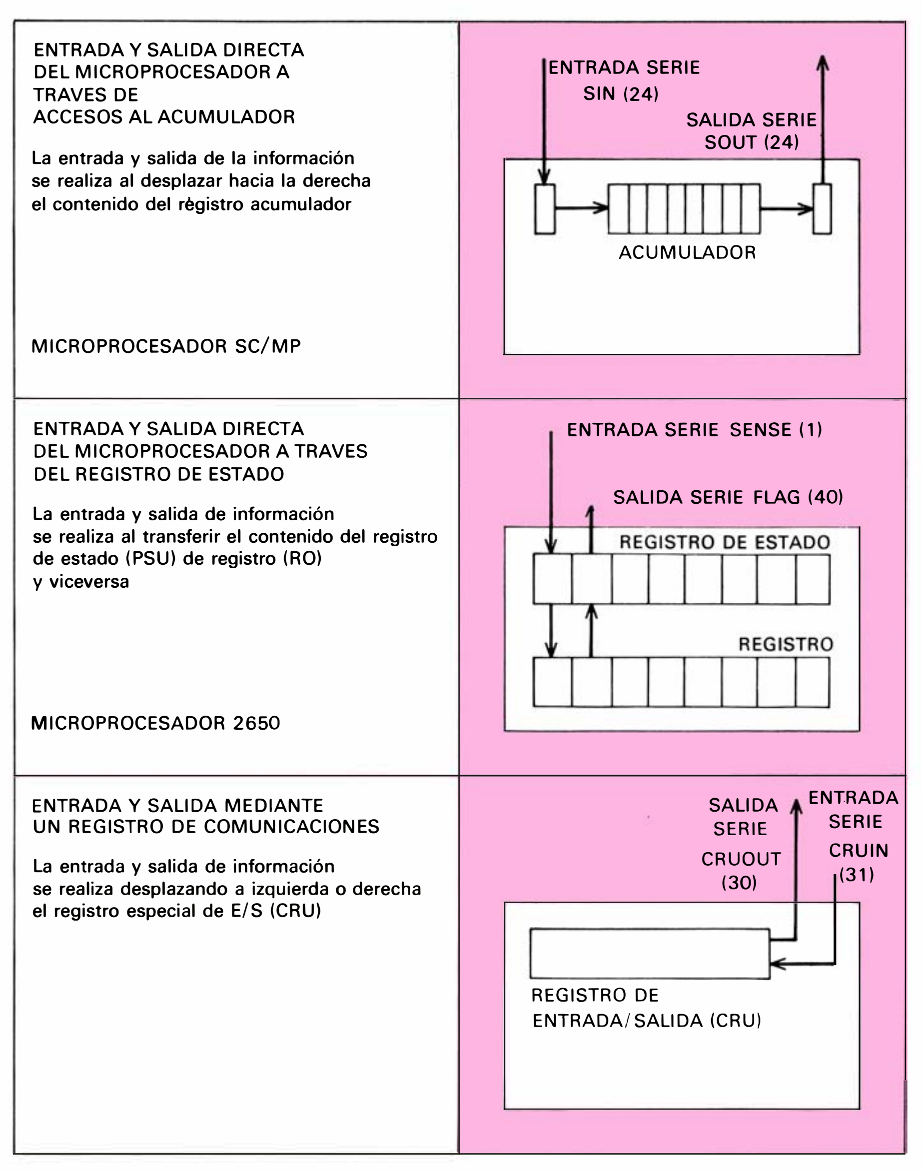
Las estructuras de información

Según los tipos distintos de funciones que realizan, podremos clasificar a los elementos que constituyen las estructuras de información en 3 grandes grupos: los registros del procesador, la memoria y los registros de entrada-salida.

Todos los fabricantes han dotado a la unidad de proceso de un conjunto de registros, que serán manejados por las instrucciones. En el número de registros y en las funciones que se les ha asignado, radica una de las principales diferencias entre los microcomputadores, con lo cual una vez más el fabricante ha querido diferenciar su máquina de las de la competencia.

Con el fin de entender las diferentes tendencias en cuanto a la organización de los registros del procesador, partiremos, consiguientemente, de los siguientes puntos fundamentales: *a)* La transferencia de información entre registros internos al microprocesador es más rápida que la transferencia de información entre registros o memorias externos, por el hecho de no tener que utilizar en el primer caso ningún mecanismo de comunicación entre pastillas. *b)* Al realizar un salto a su subprograma o atender a una interrupción, es necesario almacenar toda aquella información requerida para que, al finalizar dichas operaciones, sea posible reconstruir el estado en el que se encontraba el procesador en el momento en que se produjo la ruptura de secuencia.

En orden a la resolución de estos problemas se orientan gran parte de las soluciones adoptadas en la organización de los registros del procesador, entre las que destacaremos las siguientes, como las más significativas: en *primer* lugar, la solución más clásica consiste en dotar al procesador de un único banco de registros; en este caso, las transferencias entre dichos registros serán rápidas, mientras que salvar su contenido en memoria constituirán operaciones más lentas; en *segundo* lugar encontramos la solución consistente en duplicar los bancos de registros del procesador, controlando por programa cuál de ellos interviene en un momento dado, no siendo necesario vaciarlos en memoria al realizar un salto a subprograma; en *tercer* lugar tenemos los sistemas que disponen de un conjunto de registros con estructura de pila, lo que permitirá almacenar el contenido del contador de programa en las llamadas a subprograma; en *cuarto* lugar tenemos los sistemas en los que varios registros del procesador pueden realizar las funciones de contador de programa: en *último* lugar citaremos las máquinas en las que se ha reducido al



LA COMUNICACION DIRECTA del microprocesador con el exterior se realiza a través de los registros del procesador. En esta ilustración pueden contemplarse algunas soluciones típicas, las cuales corresponden a los siguientes sistemas: el SC/MP, el 2850, y el 9900, respectivamente.

mínimo el número de registros del procesador realizando en todas las operaciones de transferencia de información entre posiciones de memoria.

Las primitivas

Existe una fuerte relación entre la estructura de un microcomputador y su juego de instrucciones; el número y tipo de registros del procesador, el tipo de direccionado de la memoria y registros de entrada-salida, condicionarán grandemente las instrucciones que los manejan.

En principio, si bien puede decirse que cada microcomputador dispone de un juego de instrucciones propio, ocurre que todos ellos coinciden en las características de un subconjunto mínimo de instrucciones que permiten realizar operaciones muy similares, como son: las transferencias elementales de información entre registros y posiciones de memoria, las operaciones de entrada y salida de información, las operaciones aritméticas y lógicas elementales y las de salto a subprograma y retorno.

La potencia de las instrucciones de una máquina depende fundamentalmente de dos factores: de las características estructurales de la misma y de la potencia de las operaciones asociadas a un determinado código de operación.

Por lo que se refiere a las instrucciones de movimiento de datos entre registros y posiciones de memoria, sus posibilidades estarán fuertemente condicionadas por la naturaleza de los registros del procesador, lo que se traducirá fundamentalmente en las distintas formas en que puede direccionarse, desde una instrucción, el contenido de la memoria. Así como la mayor parte de las máquinas disponen de direccionado de tipo inmediato, directo y con un nivel de indirección, el direccionado indexado y el relativo al contador de programa existirán únicamente en aquellos microcomputadores dotados del soporte físico necesario para realizar este tipo de operaciones.

Además de las características estructurales, el tipo de operaciones asociadas a un código de operación influye en la potencia de las instrucciones.

Por lo que se refiere a las instrucciones de entrada-salida de información, diremos que todas las máquinas disponen de posibilidades similares. No obstante, las diferencias principales se centran en las posibilidades de comunicación en el exterior desde los registros del procesador; ello implica la aparición de instrucciones especiales de entrada y salida.

Hay otro conjunto de primitivas que están orientadas fundamentalmente a permitir la comunicación del microcomputador con otras máquinas digitales externas. Este tipo de primitivas pueden estudiarse a través de las posibles acciones entre un microcomputador y una máquina exterior.

Una máquina externa puede estar interesada en forzar al microcomputador a ejecutar una determinada función, distinta de la que está ejecutando en un momento dado, por lo que será necesario interrumpirlo; también puede estar interesada en suspender las actividades del microcomputador para evitar que progrese la ejecución de la tarea que esté ejecutando o para utilizar alguno de los recursos asignados a la máquina, como en el caso del acceso directo a memoria.

Referentes al primer tipo, diremos que cada fabricante ha dotado a su máquina de los elementos necesarios para que al producirse la petición de interrupción se proporcione al microcomputador toda la información correspondiente a la nueva actividad que se desea que se ejecute. Se han adoptado soluciones distintas; citemos algunas. En primer lugar, sistemas con una única señal de interrupción, que provoca el salto a una determinada posición de memoria. En segundo lugar, sistemas en los que la petición de interrupción desencadena la ejecución automática de una serie de operaciones destinadas a captar del exterior un código (vector de interrupción) asociado al elemento que solicitó la interrupción.

En tercer lugar, sistemas en los que el vector de interrupción se aplica directamente a las patillas de la pastilla microcomputador o microprocesador, con lo que la función de captación necesaria en la solución anterior queda simplificada, reduciéndose el tiempo de respuesta necesario para atender una petición de in-

terrupción. Por último, aquellos sistemas en los que cada posible petición de interrupción accede directamente a una patilla del circuito ocasionando el salto a una determinada dirección de la memoria.

En los tres primeros casos, se trata de soluciones en las que la información sobre la operación a realizar se da en forma codificada, mientras que, en el último, la petición de interrupción aparece ya decodificada. Hay que añadir que, en determinados sistemas, coexisten algunas de estas soluciones.

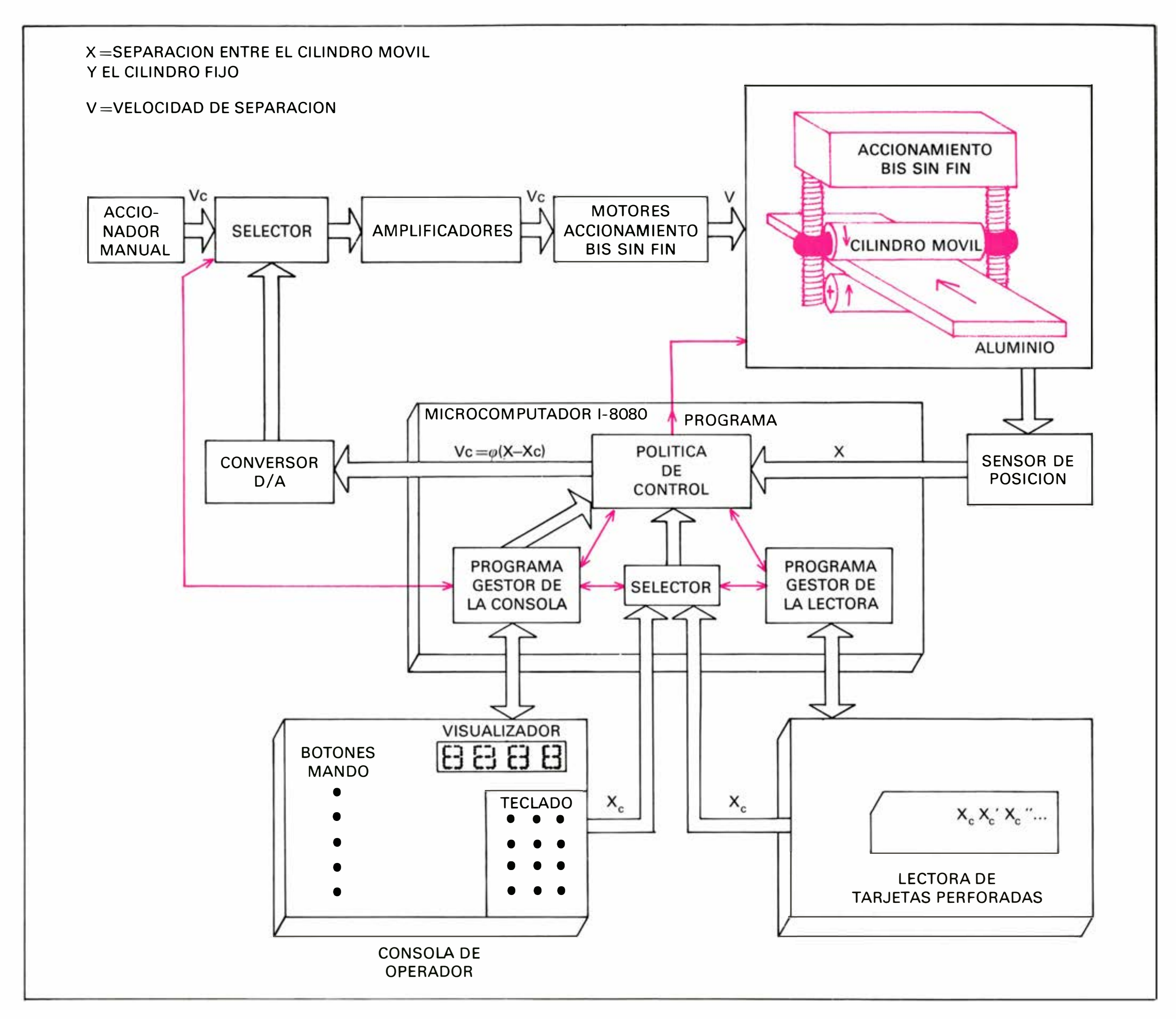
Independientemente de la forma de identificar la causa de la interrupción, en muchos microprocesadores existen mecanismos para determinar la prioridad con que deben ser atendidas las distintas interrupciones, en el caso de que entraran en conflicto.

Para finalizar, nos ocuparemos de las primitivas orientadas a provocar la suspensión de las actividades propias de la máquina.

En los sistemas basados en microprocesador la suspensión de las actividades del procesador puede tener, entre otras, las siguientes utilidades: aumentar la duración de los ciclos de lectura o escritura de memoria, con lo que se consigue poder utilizar memorias cuyo tiempo de acceso sea superior al previsto en el ciclo de ejecución de una instrucción (se trata de una suspensión momentánea de las actividades de la máquina); e inhibir el microprocesador del resto del sistema y realizar accesos directos a memoria. Casi la totalidad de las máquinas disponen de una entrada de este tipo, de forma que cuando se recibe la señal adecuada, tiene lugar la suspensión de las actividades, colocando sus patillas de acceso a las líneas de comunicación en estado de alta impedancia; lo que equivale a una desconexión electrónica de la pastilla del resto del sistema.

A modo de conclusión podemos resaltar que existen en todo microprocesador dos tipos de primitivas: aquellas cuya función básica es realizar operaciones de una sola máquina, y aquellas que son necesarias para el sincronismo y la comunicación de máquinas, trabajando paralelamente y con posibilidades de cooperación.

A nuestro entender, la tendencia en la evolución de los sistemas digitales se orienta hacia la utilización concurrente de varios microcomputadores, por lo que es de esperar la aparición de primitivas especializadas en la comunicación intermicrocomputador.



LAMINADOR DE ALUMINIO controlado por un microcomputador, basado en el microprocesador 8080. El posicionado del cilindro superior se realiza mediante una política de control que genera una señal que regula la velocidad de los motores. El microcomputador gestiona la comunicación con el operador a través de una consola y un lector de tarjetas perforadas (por cortesía de Control y Aplicaciones S.A. y el L.I.D.E.)

Un algoritmo es un conjùnto finito de órdenes cuya ejecución da lugar a una secuencia de operaciones que permiten la resolución de un problema, esto es, que partiendo de unos datos, después de un número finito de pasos, permite la obtención de los resultados.

Los algoritmos

Una máquina de utilización general será capaz de ejecutar un conjunto de órdenes tales que,en principio, cualquier algoritmo podría expresarse utilizando únicamente órdenes de dicho conjunto. Para que esto sea posible, será necesario que dichas órdenes den lugar a operaciones elementales, reducidas fundamentalmente a algunas operaciones aritméticas (adición, substracción), algunas operaciones lógicas ("y", "o", "o-exclusiva", "complemento") así como un conjunto de operaciones de manejo de la información, a nivel de los dígitos binarios. La ejecución de operaciones muy elementales, capaces de ser ejecutadas en algunas millonésimas de segundo, da la posibilidad de ejecutar un número del orden de 10^5 operaciones elementales por segundo.

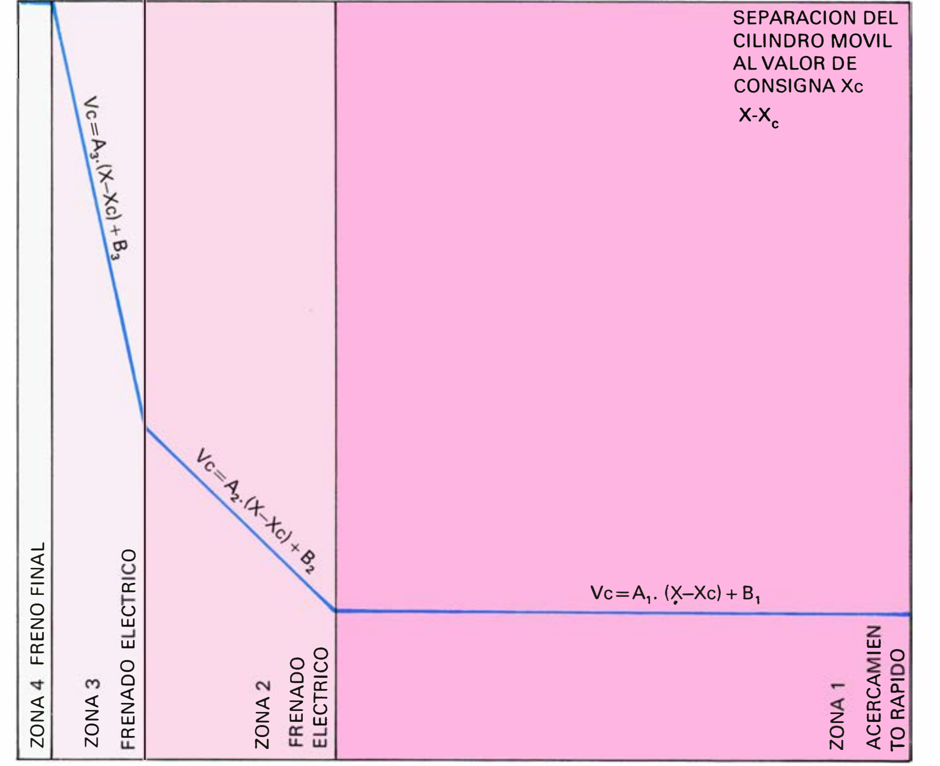
La expresión de un algoritmo –en el lenguaje de un determinado computador– se le denominará programa. Una vez elaborado, quedará almacenado en la *memoria* del computador para su posterior ejecución. Un programa quedará caracterizado fundamentalmente por dos parámetros: *a)* Su longitud, es decir, el número de instrucciones de que se compone; es lo que se traduce en la cantidad de memoria necesaria para almacenar dicho programa. *b)* El volumen de datos que maneja, lo cual determina la cantidad de memoria necesaria para su almacenamiento.

Así pues, el volumen de elementos de memorización de la información, capaces de ser manejados por un computador, impondrán una primera limitación al tipo de algoritmos que puedan ser realizados en una determinada máquina.

Por otro lado, la realización de las operaciones elementales toma cierto tiempo, por pequeño que éste sea y, en consecuencia, todo programa estará caracterizado por un tiempo de ejecución, lo que dará lugar a una nueva limitación en el tipo de algoritmos que podrán ser resueltos en una máquina particular.

Para determinar si una máquina puede soportar un algoritmo es necesario responder a las siguientes cuestiones: El número de elementos de memorización ¿es suficiente para poder almacenar los datos y el programa del algoritmo? El tiempo de ejecución del algoritmo, en la máquina propuesta, ¿es adecuado para la aplicación considerada?

Es interesante resaltar que algunos algoritmos no pertenecen a la clase *A* del triple que define funcionalmente la máquina *(D, P, A)* por limitaciones de espacio de memorización, otros por no satisfacer la restricción temporal o bien ambas a la vez, constatándose una evolución tecnológica que permite ampliar la clase *A* en las nuevas máquinas que aparecen con nuevas tecnologías más rápidas de proceso de la información (menor tiempo para ejecutar las operaciones elementales) y con capacidades de memorización mayores para la misma superficie de material semiconductor.



POLITICA DE CONTROL que da la velocidad de consigna de los motores de acercamiento de los cilindros, en función de la distancia que falta por recorrer hasta obtener el espesor deseado. La política queda definida por 9 parámetros (3 de delimitación de zona y 2 por segmento rectilíneo).

Una aplicación

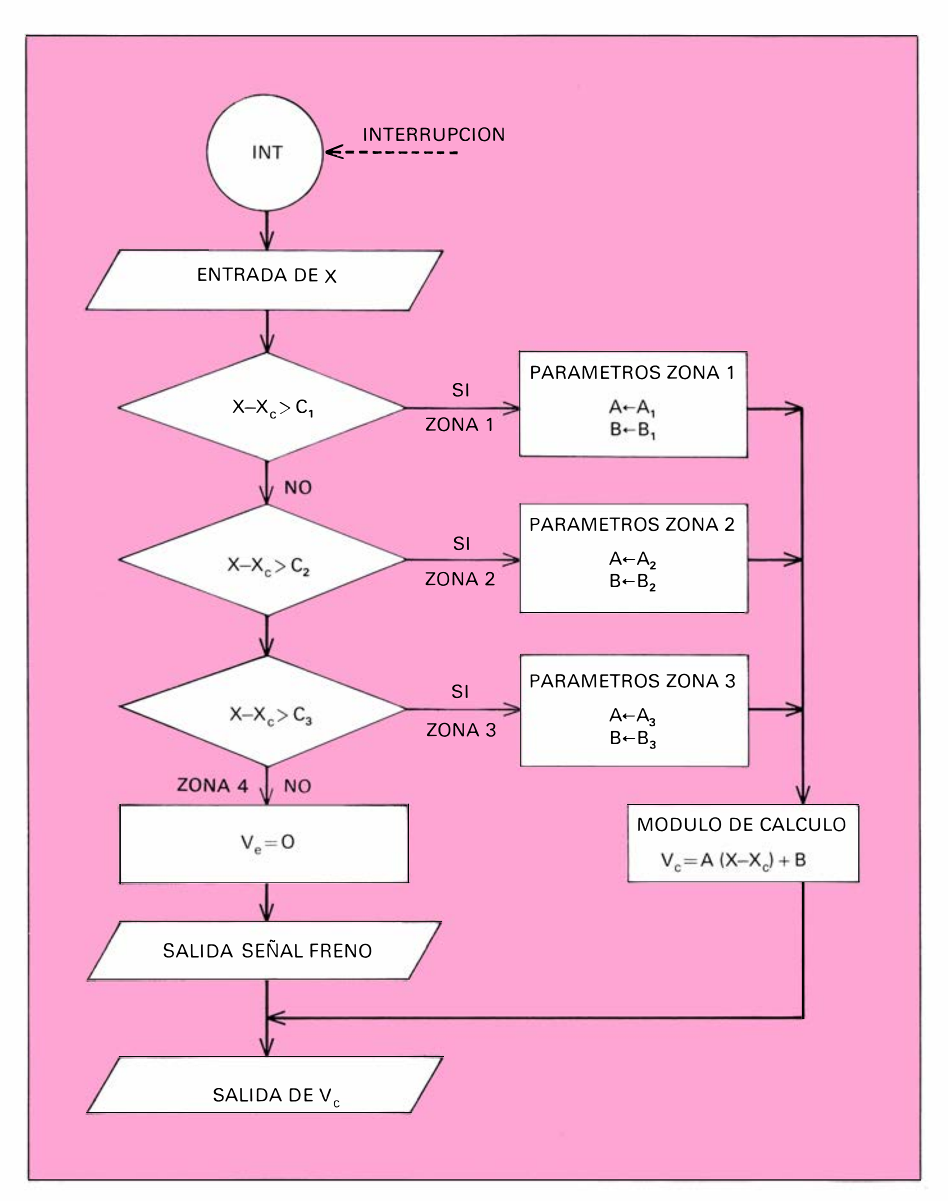
Veamos un caso-estudio de un sistema de control en tiempo real basado en un microprocesador 8080, cuyo algoritmo de control tuvo que ajustarse a las restricciones aquí consideradas. Por un lado, la "restricción temporal", de forma que la ejecución no podía rebasar un tiempo de cálculo del orden de la milésima de segundo y, por otro, tuvo que ajustarse a la "limitación de memoria", que para el microprocesador utilizado no podía rebasar un máximo de 64 K palabras de 8 bits (1 K=1024); ello desaconsejaba la realización de algoritmos de control del tipo acceso a tablas que exigen gran cantidad de memoria.

Se trata de un proceso de laminado de aluminio en caliente en el cual se efectúan una serie de pasadas de una placa de dicho material entre dos cilindros giratorios, cuya separación se reduce en cada nueva pasada. De los dos cilindros, uno es fijo y el otro se desplaza verticalmente gracias a la acción de unos motores. Este posicionado del cilindro móvil debe ser rápido y preciso, con el fin de optimizar el proceso de laminado.

La política de control consiste básicamente en dar una velocidad de acercamiento a la posición de consigna del cilindro móvil, según una ley no lineal formada por 4 segmentos rectilíneos.

Al producirse una entrada en el microcomputador del valor de posición del cilindro móvil *X* se determina primero la zona que corresponde al valor leído. A continuación se generan los parámetros correspondientes a la zona, pasándose seguidamente al módulo de cálculo de la velocidad de consigna *Vc* de los motores de accionamiento del cilindro móvil. Cuando el cilindro móvil ha llegado ya a décimas de milímetro de la posición deseada se da por finalizada la operación de posicionado, actuándose un freno mecánico.

Es interesante constatar que, dada la limitación temporal, para realizar el

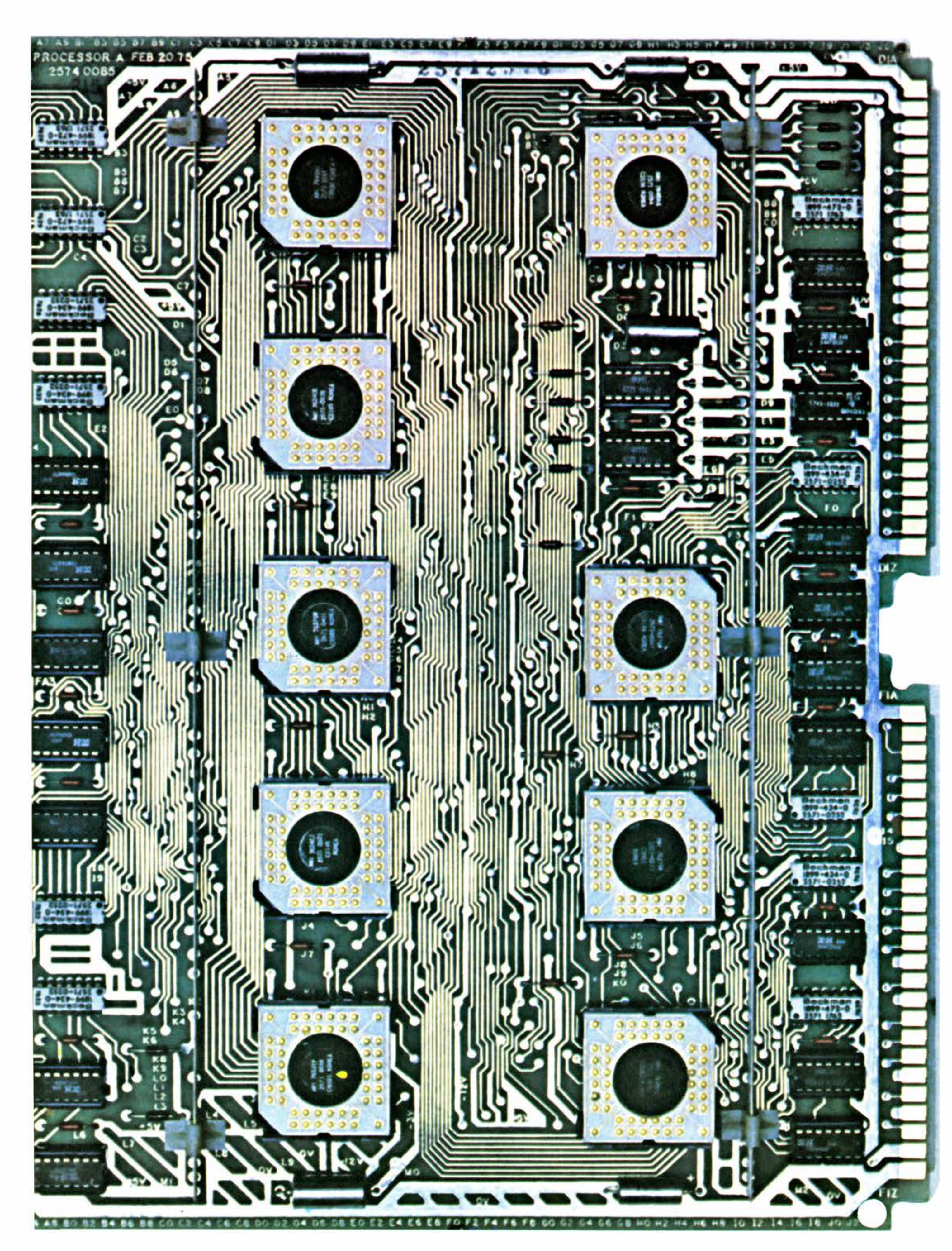


ALGORITMO de control. Permite generar la velocidad de consigna según la política de control que determina la zona en la que está el cilindro móvil y calcula el valor de la velocidad correspondiente. Al llegar a la cuarta zona (posición dentro de la tolerancia permitida), se da la orden de frenado.

control de los motores se tuvo que recurrir a una política de control sencilla y con poca exigencia de cálculos (bastaron una suma, una resta y una multiplicación) para cada nuevo valor de consigna generado. Por otra parte, la política es muy fácil de almacenar, pues queda acotada con los nueve parámetros que definen los segmentos rectilíneos. El programa total que incluye a su vez un programa gestor de una lectora de tarjetas perforadas, otro programa gestor de la consola del operador, un programa gestor de interrupciones y un programa gestor del estado global del autómata, además del programa del algoritmo de control que acabamos de ver sólo ocupa 3 K palabras de 8 bits, por lo que se satisface holgadamente la limitación de memoria que, como máximo en este microcomputador, puede llegar a 64 K. En consecuencia, la máquina digital es aún expandible, y en los momentos que se escriben estas líneas se está estudiando la posibilidad de añadir ciertas funciones adicionales al microcomputador.

Cabe reflexionar sobre las alternativas clásicas de haber diseñado la máquina a base de componentes MSI. o SSI o incluso utilizando exclusivamente técnicas analógicas. No creemos que sería aventurado afirmar que el sistema basado en el módulo de utilización general LSI (microprocesador), es más fiable, más económico que cualquier otra alternativa con las mismas prestaciones y más adaptable y ampliable incluso después de instalado.

Esta última característica permitió diseñar la política de control de un sistema, que conocíamos sólo cualitativamente, ajustando los parámetros en el momento de la instalación y, por otra parte, satisfacer nuevas necesidades incorporando nuevas funciones a medida que los usuarios las iban descubriendo.


PROCESSOR A FEB 20 75
2574 0085

Microelectrónica y proceso de datos

Los grandes ordenadores de nuestros días no podrían existir sin los componentes microelectrónicos. La microelectrónica ha provocado también la aparición de otros más pequeños, capaces de realizar tareas en que no es necesario utilizar los grandes

Lewis M. Terman

Los modernos ordenadores y los dispositivos microelectrónicos se han interrelacionado de un modo tan estrecho en su evolución que pueden considerarse virtualmente simbióticos. Los primeros ordenadores digitales, de gran dimensión y de programas almacenados, estaban basados en tubos de vacío; sus sucesores, en transistores discretos. Tales elementos no podrían usarse para construir ordenadores dotados con las capacidades de los que hoy en día poseemos. Resultarían excesivamente costosos y poco fiables; además, requerirían demasiada energía. Serían, sobre todo, desmesuradamente grandes y lentos. El número de operaciones por segundo que puede realizar un ordenador depende de la velocidad con que sus elementos activos se comunican entre sí, y esa velocidad depende de la rapidez con que operen los elementos activos y de la proximidad a que se hallen. La tecnología de los ordenadores ha urgido la microelectrónica y ésta ha evolucionado ampliamente como respuesta a tal exigencia.

Los mismos ordenadores han evolucionado, en gran parte, en respuesta a la demanda de proceso de datos: grandes cantidades de caracteres, sean números o letras, que representan información de casi todos los tipos imaginables. En actividades científicas, los ordenadores recopilan y analizan un gran número de mediciones, como las posiciones y las velocidades secuenciales de una astronave, y resuelven problemas matemáticos extraordinariamente largos y complejos, como la trayectoria de esa misma astronave. En actividades comerciales, registran y procesan inventarios, compras, facturas, nóminas, depósitos bancarios, etc., y siguen la pista de las transacciones comerciales en curso. En la industria, supervisan y controlan los procesos de fabricación. En la esfera de gobierno, elaboran estadísticas, analizan información económica y calculan impuestos. Hoy sería difícil hallar una tarea que requiera el procesamiento de grandes cantidades de información que no la lleven a cabo los ordenadores. Más aún, los ordenadores realizan también numerosas tareas de menor envergadura, que exigen cantidades comparativamente más pequeñas de procesamiento de datos.

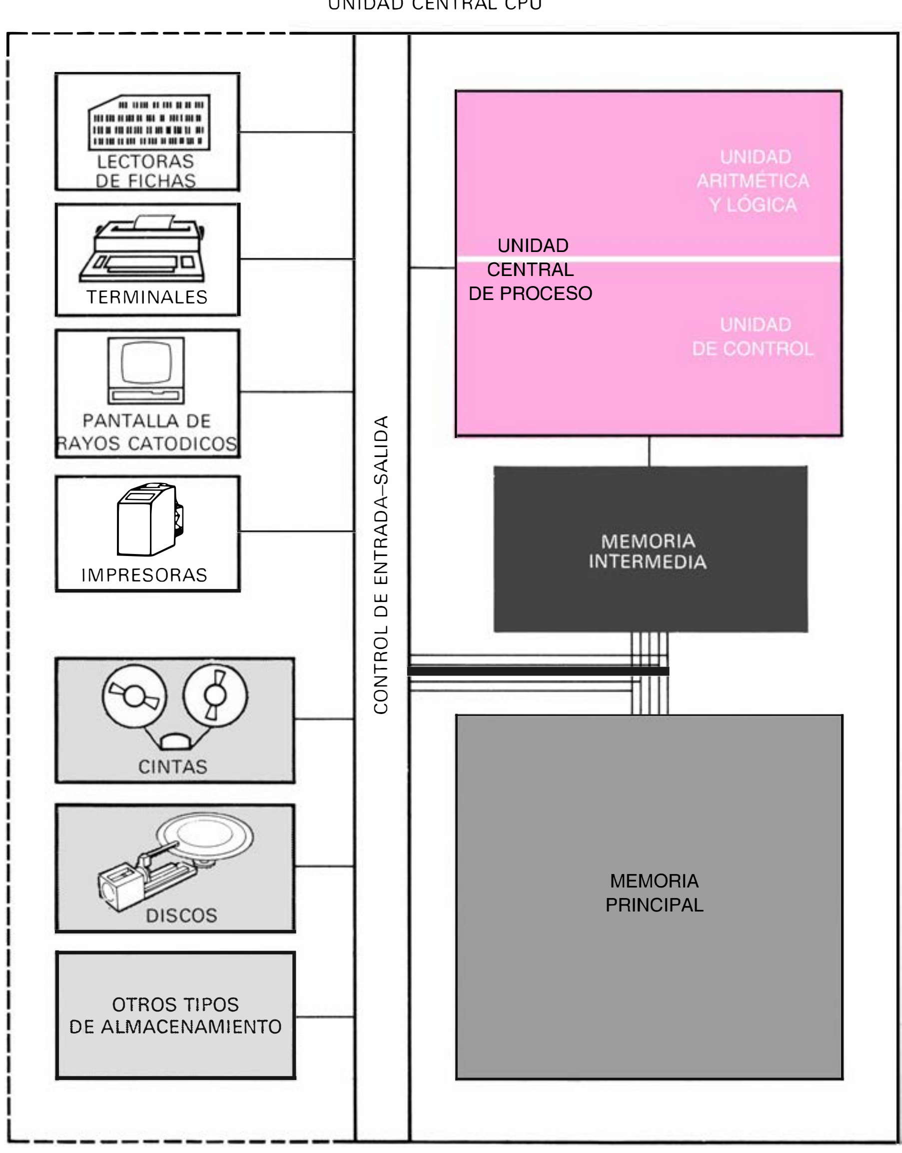
Si bien un gran sistema de proceso de datos es costoso, puede efectuar millones de operaciones por segundo; incluso los sistemas pequeños pueden realizar 100.000 operaciones por segundo. En esto reside el motivo fundamental del crecimiento del proceso de datos por ordenador: un ordenador puede efectuar tantas operaciones en tan poco tiempo que el coste de cada una resulta muy bajo.

El elemento clave del crecimiento del proceso de datos por ordenador es, quizás, el hecho de que, con la aparición de la microelectrónica, el coste de los componentes físicos del ordenador ha disminuido constantemente, mientras la capacidad y el rendimiento han aumentado ininterrumpidamente. El coste de un ordenador dividido por su potencia de cálculo (medida en millones de operaciones por segundo) es hoy más de 100 veces menor que en 1960. En los mismos años, el tiempo de ciclo (tiempo necesario para efectuar una operación) de las máquinas más grandes se ha hecho 10 veces menor. El aumento del rendimiento de los elementos físicos del ordenador ha revestido importancia para el desarrollo de los ordenadores centrales o de "unidad principal": las máquinas mayores y más rápidas. Ha sido aun más importante para el desarrollo de los ordenadores periféricos: las máquinas más lentas y pequeñas. Numerosas tareas no requieren la gran capacidad, flexibilidad y velocidad de proceso que es propia de un

TABLERO DEL PROCESADOR de un pequeño ordenador de "unidad principal" Burroughs B-80, que contiene la lógica de decisión de todo el ordenador: es responsable de la ejecución de los programas introducidos en la máquina. Las medidas reales del tablero del procesador son de 25 × 32 centímetros. Los componentes lógicos son nueve pastillas que contienen circuitos integrados a gran escala, de óxido metálico semiconductor (MOS), alojada cada una de ellas en uno de los nueve paquetes blanco y oro. En este tablero, las funciones lógicas, en vez de ejecutarse mediante circuitos lógicos, están almacenadas en forma de configuraciones permanentes de bits en pequeñas memorias de sólo lectura, incluidas en los paquetes blanco y oro superior, medio e inferior de la columna izquierda. Los paquetes segundo y cuarto de esa columna alojan las pastillas de registros de trabajo, que almacena datos para su ejecución en el procesador. La pastilla superior de la columna derecha controla el funcionamiento de los dispositivos de entrada-salida del ordenador. La segunda pastilla de esa columna es el registro de direcciones de la memoria principal, que contiene la dirección de la palabra de bits a la que se está accediendo desde la memoria principal del ordenador. La pastilla más cercana a la parte inferior de la columna derecha es una micropila, otro tipo de registro de direcciones de memoria. En la parte inferior de esa columna, se encuentra la pastilla de "sincronización-estado de máquina", que controla la sincronización y secuencia de las operaciones ejecutadas por todo el tablero del procesador. Los demás componentes del cuadro no tienen funciones lógicas. Los rectángulos negros son amplificadores. Los rectángulos blancos y los pequeños cilindros rayados son resistencias. Los componentes en forma de bombonas, los cilindros y los grandes cilindros plateados son condensadores. Las dos barras blancas que atraviesan el cuadro en su totalidad son barras colectoras que distribuyen los cables de electricidad y tierra que recorren el cuadro. Todos los componentes están interconectados mediante estructuras estañosoldadas. El tablero se inserta en el ordenador a través de conectores de oro, sitos a lo largo de sus bordes.

sistema de unidad principal, aunque sea pequeño. En los años 60, los avances logrados en componentes microelectrónicos condujeron al desarrollo del miniordenador, seguido posteriormente por el microordenador, que es aun menor. Ambos han satisfecho la necesidad de sistemas de proceso pequeños, pero relativamente flexibles, capaces de ejecutar funciones de cálculo comparativamente sencillas a menor coste. Además, con el desarrollo del microprocesador, se ha producido una división casi darwiniana de los ordenadores en diferentes especies, surgiendo máquinas de diferentes tamaños y organizaciones, adaptada cada una de ellas a una gama diferente de funciones. El progreso hacia ordenadores más pequeños tiene grandes posibilidades de continuar; se habla ya de nanoordenadores y picoordenadores.

Actualmente, un gran sistema de ordenador central puede estar formado por 100.000 circuitos lógicos y entre cuatro y ocho millones de octetos de memoria. (Un octeto consta de ocho o nueve bits; un bit es un dígito binario.) Un sistema central pequeño podría tener 20.000 circuitos lógicos y 128.000 octetos de memoria. Un sistema típico de miniordenador tendría entre 5000 y 10.000 circuitos lógicos y una memoria de 16.000 a 32.000 octetos. Un sistema de microordenador podría tener de unos 1000 a 2000 circuitos lógicos y de varios centenares a varios millares de octetos de memoria. Existe una correspondencia en cuanto al coste de dichos sistemas: los ordenadores centrales mayores cuestan varios millones de dólares y un sistema completo de microordenador apenas alcanza la cifra de unos centenares de dólares.



LOS SISTEMAS DE ORDENADOR DE UNIDAD PRINCIPAL u ordenadores centrales concentran la mayor parte de la capacidad del proceso de datos del ordenador en la unidad central de proceso [CPU, del inglés "central processing unit"] *(color).* **Físicamente, la estructura esencial de la CPU de un gran sistema de ordenador incluye la unidad central de proceso propiamente dicha, la memoria intermedia de alta velocidad y la memoria principal, y puede incluir también varias vías de entrada-salida. La unidad central de proceso propiamente dicha consta de una unidad que ejecuta las operaciones aritméticas y lógicas del programa del ordenador y otra unidad que controla el funcionamiento de todo el sistema ordenador. El programa de ordenador y los datos en uso activo se almacenan en la memoria intermedia** *(gris oscuro),* **a la cual puede acceder la unidad central de proceso, a alta velocidad. Además de la memoria intermedia, hay una memoria principal mayor y más lenta** *(gris medio).* **Alejados de la estructura esencial de la CPU, están los discos, cintas y otros sistemas de almacenamiento-memoria** *(gris claro),* **que son los más grandes y lentos. Los dispositivos de entrada-salida son lectoras de cinta, lectoras de fichas, teclados, pantallas de tubos de rayos catódicos, discos, unidades de cinta, teleescritores y otros.**

La mayoría de los sistemas de ordenador, independientemente de su tamaño, constan de tres elementos básicos: los accesos o vías de entrada salida, la jerarquía de memorias y la unidad central de proceso. Las vías de entrada salida son aquellos caminos a través de los cuales la información (instrucciones y datos) entra y sale del ordenador por medios diversos, como fichas perforadas, cintas magnéticas y terminales. La jerarquía de memorias almacena las instrucciones (el programa) y los datos en el sistema, de modo que puedan recuperarse rápidamente a petición de la unidad central de proceso. Esta unidad central de proceso, que consta de una unidad lógica de control y una unidad aritmética y lógica, dirige el funcionamiento de todo el sistema por la emisión de mandatos a otras partes del sistema y por su actuación sobre las respuestas. Lee información de la memoria, interpreta instrucciones, realiza operaciones con los datos de acuerdo con las instrucciones, graba otra vez los resultados en la memoria y transfiere información entre los niveles de memoria o a través de las vías de entrada/salida. Las operaciones que realiza con los datos pueden ser o aritméticas, como sumas o restas, o lógicas, como por ejemplo la comparación de dos campos de números para determinar si son o no equivalentes.

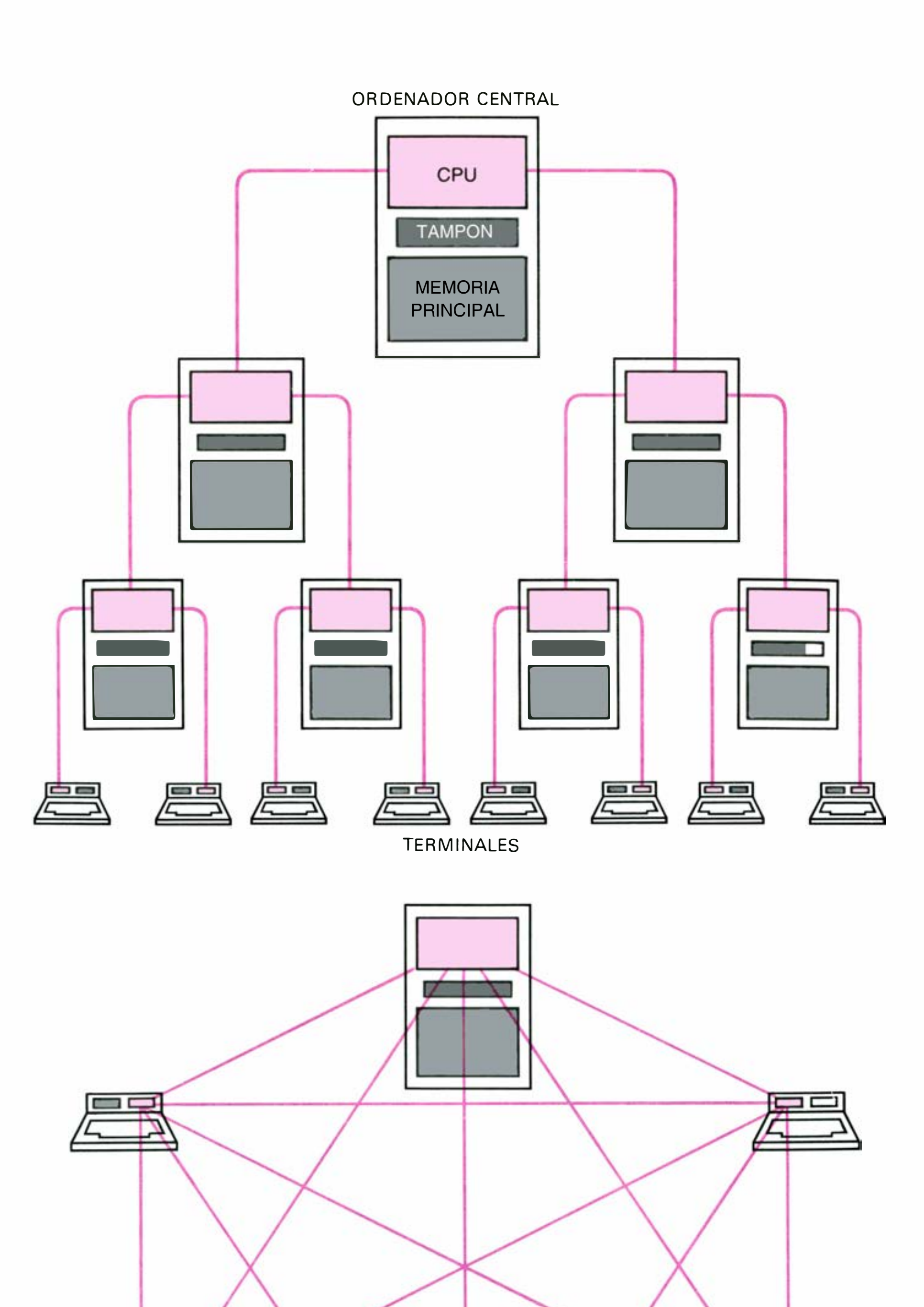
La organización de la memoria de un ordenador es jerárquica tanto en términos de velocidad como de coste. En interacción directa con la unidad central de proceso, hay una memoria intermedia que transfiere información, a alta velocidad, a la unidad central de proceso. Sin embargo, la velocidad es cara y el coste de cada bit de información almacenado en esta memoria intermedia es alto. Inmediatamente debajo de esta memoria intermedia, en la jerarquía, está la memoria principal del ordenador, que es mayor y más lenta que la memoria intermedia citada y cuyo coste por bit de información almacenado es mucho menor. Por debajo de la memoria prin-

cipal, aparecen los sistemas de almacenamiento, como discos y cintas magnéticas, mayores y más lentos y cuyo coste por bit es todavía menor. Así pues, los niveles de memoria más alejados de la unidad central de proceso son progresivamente mayores, más lentos y baratos.

La información que actualmente está en uso activo se transfiere a la memoria más rápida (la memoria intermedia) para que sea más fácil acceder a ella; la información de menor actividad o inactiva se coloca en el almacenamiento de menor coste por bit. De este modo, la unidad central de proceso encuentra en la memoria intermedia la información que requiere para procesarla inmediatamente; sólo ocasionalmente ha de recuperar información situada en niveles inferiores de la jerarquía. El rendimiento de todo el sistema del ordenador está determinado por la velocidad de la memoria intermedia, en tanto que la capacidad del sistema y el coste por bit quedan determinados por el número y la dimensión de los niveles de almacenamiento.

El impacto de la microelectrónica en las memorias de los ordenadores ha sido diferente en los distintos niveles de la jerarquía. Desde el principio, la memoria intermedia ha constado solamente de circuitos electrónicos. Con la aparición de la microelectrónica, los circuitos se han hecho menores y más rápidos que sus predecesores de hace cinco o diez años y han contribuido a aumentar la velocidad de los ordenadores centrales.

Los dispositivos microelectrónicos han transformado la estructura de la memoria principal en un grado todavía mayor. De la mitad de los años 50, atravesando la década de 1960, los elementos individuales de almacenamiento de información en la memoria principal de un ordenador eran los "núcleos" magnéticos, pequeñas bobinas toroidales de ferrita en que se almacenaba un bit de información según la dirección de imantación, en torno a la bobina toroidal. Cada núcleo se fabricaba individualmente y millones de ellos se agrupaban en grandes series mediante cables que pasaban por el orificio central de cada uno de ellos. La información se introducía en los núcleos y se extraía de ellos por lectura mediante circuitos electrónicos de soporte. Debido a las características de los núcleos y al débil acoplamiento magnético entre los núcleos y los cables en que se agrupaban, se requería una compleja red de circuitos de soporte para escribir y leer la información en los núcleos y no era posible fabricar los circuitos ni los núcleos como circuitos integrados.

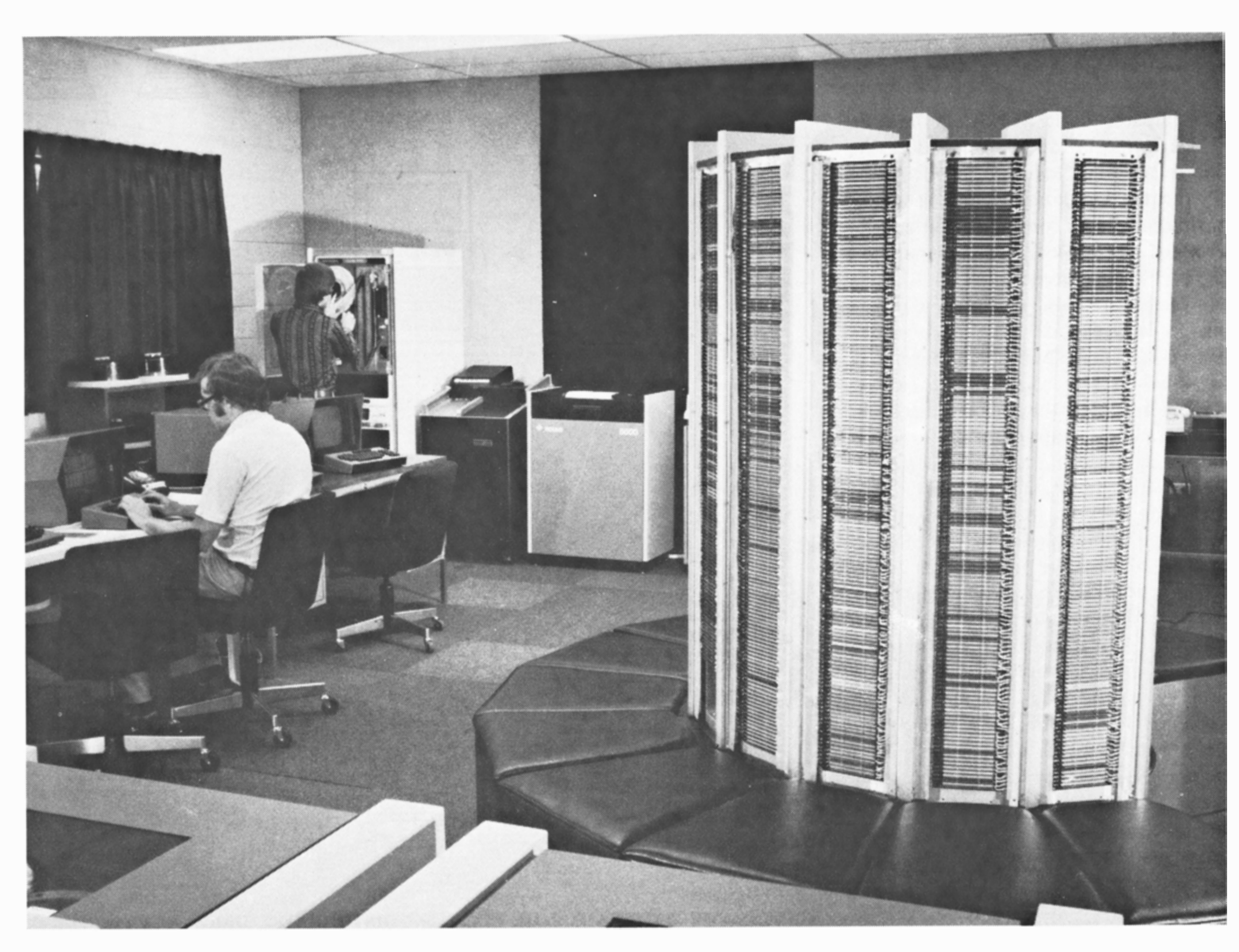


SISTEMA DE ORDENADOR DE PROCESO DISTRIBUIDO. Distribuye la capacidad de proceso a través de una red de ordenadores, en lugar de concentrarla en una unidad central de proceso. Físicamente, esto permite a la red extenderse por una amplia zona geográfica y hace posible que cada procesador local *(pequeños recuadros en color)* actúen sobre la información independientemente, antes de transmitirla a uno o más elementos de la red. Las dos formas básicas de un sistema de proceso distribuido son la estructura jerárquica *(arriba)* y la estructura de igualdad *(abajo)*. En la estructura jerárquica, hay un ordenador central que comunica con ordenadores menores situados en un nivel inferior, cada uno de los cuales se comunica con ordenadores de un nivel inmediatamente inferior, y así sucesivamente. En la estructura de igualdad, todos los elementos están en el mismo nivel y cada elemento puede comunicarse con los demás. En la práctica, ambas formas se combinan en un sistema híbrido. A la memoria intermedia del texto denominamos aquí tampón.

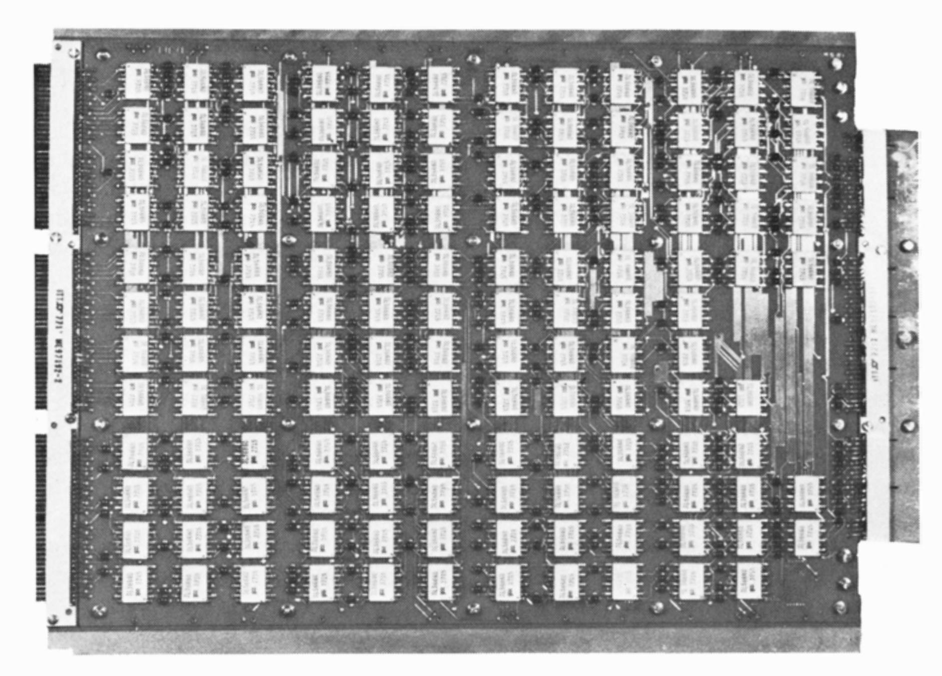
A principios de los años 70, se desarrollaron celdas de memoria semiconductoras que servían para la misma finalidad que los núcleos, y comenzaron a instalarse circuitos de memoria integrados como memoria principal de los ordenadores. En un circuito de memoria integrado, el mecanismo de almacenamiento de información (la celda de memoria) no es una pieza pasiva de material, sino que es por sí mismo un circuito semiconductor. Además, los circuitos de soporte necesarios para escribir información en la celda de memoria y extraerla de ella por lectura son mucho más sencillos que en el caso de los núcleos. Así, tanto la celda de memoria como los circuitos de soporte pudieron combinarse en la misma pastilla de silicio. Los dispositivos semiconductores microelectrónicos aparecen en la memoria principal de la inmensa mayoría de los ordenadores que se producen hoy, aunque siguen fabricándose núcleos magnéticos para ciertos fines.

Una característica especialmente valiosa de una memoria principal de ordenador, formada por dispositivos microelectrónicos, es que el coste por bit de memoria es prácticamente independiente del tamaño de la memoria. Con las memorias de núcleos, la red de circuitos de soporte era compleja y cara, fuese cual fuere el tamaño del sistema, hecho que hacía que, en un ordenador con una memoria de núcleos pequeña, el coste por bit fuera mucho más alto que en un ordenador con una memoria de núcleo grande. Con las memorias microelectrónicas, el coste del sistema depende principalmente de las pastillas de la memoria y de su encapsulado, no de los circuitos de soporte. De ahí, que el coste de un sistema de ordenador sólo sea proporcional al tamaño de la memoria y que el coste por bit de memoria permanezca fijo. En otras palabras, el usuario de un sistema de miniordenador o microordenador subordinado que sólo necesite una pequeña cantidad de memoria, no tiene ya que pagar una importante penalización en el coste por bit del sistema. Este factor ha tenido una incidencia decisiva en el crecimiento de los pequeños sistemas de ordenador subordinados.

La microelectrónica no ha llegado muy lejos en la construcción de dispositivos de almacenamiento, como discos o cintas magnéticas, y probablemente no avanzará demasiado en un futuro previsible (véase "Memorias microelectrónicas", de David A. Hodges, en este mismo número). Hay varias razones para ello. En primer lugar, las cintas y discos no son caros. Sin embargo, la



SUPERORDENADOR CRAY-1, fabricado por Cray Research, Inc., de Minneapolis, Minnesota. Se trata de un potente ordenador digital, científico, de alta velocidad, para la ejecución de tareas muy complejas. La unidad principal, CPU, del CRAY-1 (la estructura cilíndrica situada en el centro de la sala) se compone de más de 1000 tableros lógicos modulares, insertados horizontalmente en 24 chasis verticales. Esta unidad principal contiene la unidad central de proceso propiamente dicha y la memoria del ordenador. Los bancos situados alrededor de la base de la unidad principal son armarios que alojan los 12 alimentadores eléctricos del ordenador. A la izquierda, un operador está sentado ante el teclado del terminal de un miniordenador que funciona como unidad de control de mantenimiento del ordenador CRAY-1. Más allá, otro operador carga una cinta magnética en una unidad de cinta de la unidad de control de mantenimiento. Los armarios que aparecen en primer plano son dispositivos de almacenamiento en discos. Este ordenador CRAY-1 se halla en el laboratorio de la compañía en Chippewa Falls.



CUADRO LOGICO DEL SUPERORDENADOR CRAY-1, uno de los 1056 cuadros semejantes de la unidad central de proceso. Cada cuadro mide 15 por 20 centímetros. En él pueden montarse hasta 288 pastillas de circuitos integrados, 144 en cada lado; en el lado del tablero que vemos aquí, hay 137 pastillas. Toda la CPU del CRAY-1 está formada por sólo seis tipos diferentes de elementos: cuatro tipos de circuitos integrados *(grandes paquetes blancos)* **y dos tipos de resistencias.**

maquinaria eléctrica y mecánica necesaria para introducir y extraer la información contenida en ellos es muy costosa. Por tanto, el coste por bit de dichos sistemas de almacenamiento está en proporción fuertemente inversa al tamaño del sistema, y el menor coste por bit se obtiene con los más grandes sistemas. Por ejemplo, un sistema de almacenamiento en discos capaz de contener 100 millones de octetos de memoria sólo cuesta unas milésimas de centavo de dólar por bit. Para un gran sistema de ordenador, el coste por bit de los métodos de almacenamiento microelectrónicos nunca podría competir con el coste por bit de almacenamiento mediante disco o cinta.

Segundo, los discos y las cintas están diseñados con vistas a la estabilidad de la información: si se produce un fallo en la energía eléctrica del sistema, la información almacenada en disco o cinta permanece intacta. La información almacenada en una memoria microelectrónica se perdería al producirse un fallo de corriente. Tercero, las cintas y los discos son relativamente compactos y, en algunos casos, es importante disponer de la posibilidad física de transportarlos o almacenarlos fuera del ordenador.

En un sistema de ordenador pequeño, los requisitos son muy distintos. Por ejemplo, el usuario de un miniordenador no necesita en absoluto una capacidad similar a la de una gran unidad de almacenamiento en discos y, ciertamente, no querrá pagar su coste. Le basta con obtener un sistema menor con un coste total más bajo, incluso si el coste por bit es mayor. Por tanto, la microelectrónica puede penetrar en el área de los dispositivos de almacenamiento de menor capacidad. Los mecanismos de almacenamiento semiconductores, como son los dispositivos acoplados por corriente, llegarán a ser más económicos que las tecnologías de memoria principal de semiconductores y su coste se hará competitivo con el de los pequeños sistemas de almacenamiento en discos o cinta e incluso menor que éste. Puesto que el acceso a un sistema de almacenamiento por semiconductores sería electrónico y no físico, como lo es necesariamente en el caso de los discos y cintas, el tiempo necesario para introducir o recuperar información en un sistema de almacenamiento por semiconductores sería mucho menor. Aun cuando el sistema de almacenamiento por semiconductores cueste más por unidad, el usuario considerará aconsejable aceptar un mayor coste por un mayor rendimiento. Es todavía demasiado pronto para determinar hasta qué punto pueden llegar a ser satisfactorios los dispositivos microelectrónicos de almacenamiento.

En la jerarquía de memorias y almacenamientos de un ordenador, la información se almacena pasivamente y se recupera inalterada. En la unidad central de proceso del ordenador, se opera activamente sobre la información. Los elementos de que están formadas las partes del proceso de datos del sistema de ordenador son los circuitos lógicos. Los circuitos lógicos son claramente diferentes de los circuitos de memoria. En una memoria microelectrónica, cada celda de memoria contiene un bit de información. En un momento dado, sólo se accede a una pequeña fracción de los bits del sistema de memoria, de modo que la energía auxiliar o de reserva necesaria para mantener un funcionamiento adecuado de la memoria puede ser baja. La energía que los circuitos de soporte necesitan para escribir información en la memoria o extraerla de ésta se suministra a petición, mediante circuitos si-

			CRAY-1
UNIDAD CENTRAL DE PROCESO	NUMERO DE PASTILLAS LOGICAS		278.000
	TIEMPO DE CICLO DE CPU		12,5 NANOSEGUNDOS
MEMORIA	INTER-MEDIA	TECNOLOGIA	—
		CAPACIDAD	—
		TIEMPO DE CICLO	—
	PRINCI-PAL	TECNOLOGIA	SEMICONDUCTOR BIPOLAR
		CAPACIDAD	HASTA 1.048.576 64 PALABRAS BIT
		TIEMPO DE CICLO	50 NANOSEGUNDOS
OTROS	TAMAÑO DE CPU		3-METRO DIAMETRO BASE 1,5-METRO DIAMETRO PARTE CENTRAL 2 METROS DE ALTURA
	PESO DE CPU		5,25 TONELADAS
	PRECIO DE COMPRA		8.000.000 DOLARES

ESTA TABLA DE INFORMACION presenta varias estadísticas significativas sobre el ordenador CRAY-1. El tiempo de ciclo de la CPU es la cadencia de impulsos del reloj de la CPU. El CRAY-1 sólo tiene una memoria. Esta se divide en 16 bancos de operación independiente a los que puede accederse uno tras otro, de modo intercalado, a intervalos de 12,5 nanosegundos. El tiempo de ciclo de memoria de 50 nanosegundos es el tiempo necesario para acceder a una palabra de información de un banco; puesto que las palabras situadas consecutivamente pueden almacenarse y recuperarse de bancos sucesivos en un cuarto de ese tiempo, el tiempo de ciclo efectivo de la memoria es de 12,5 nanosegundos. El número de pastillas de la CPU es una estimación aproximada sobre el tamaño de un ordenador, ya que cada pastilla puede tener muchos o pocos circuitos lógicos.

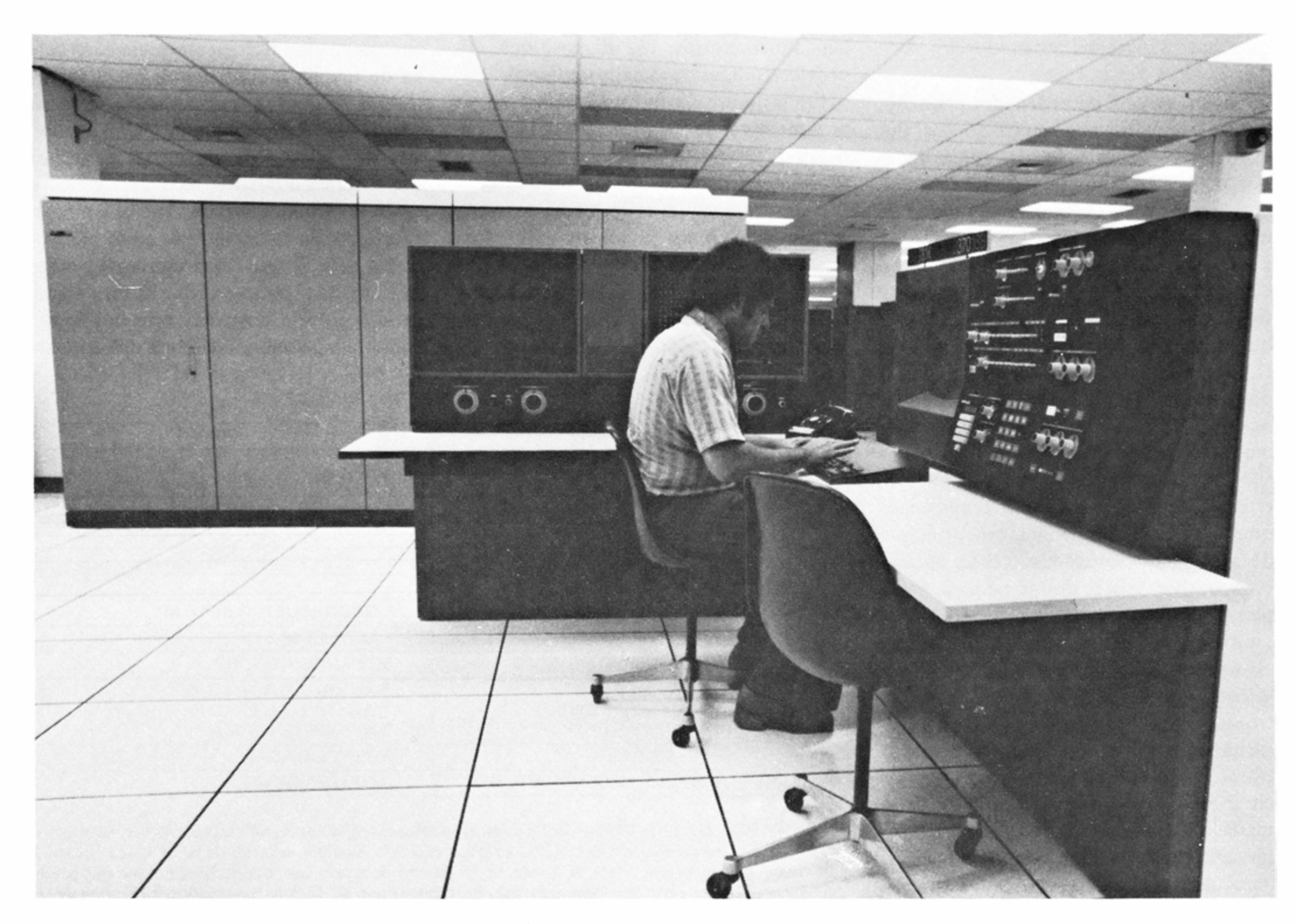
tuados fuera de la serie de celdas de memoria. En cambio, en el procesador de un ordenador, todos los circuitos lógicos están en continuo funcionamiento. Sólo unos pocos circuitos pueden estar en un instante determinado en proceso de conmutación pero todos los demás circuitos estarán dando una salida correspondiente a su entrada y habrán de estar preparados para conmutar en cualquier momento si la entrada debiera cambiar.

Un circuito lógico consta esencialmente de un dispositivo de carga con impedancia comparativamente alta (como una resistencia) conectado en serie con uno o más dispositivos de conmutación que se conectan en paralelo. Cada dispositivo de conmutación tiene una impedancia que depende de la fuerza de la tensión de entrada. Cuando ésta es alta, la impedancia del dispositivo de conmutación es comparativamente baja; la corriente fluye a través del conmutador y, de esta forma, está en actividad (*on*). Cuando la tensión de entrada es baja, la impedancia del dispositivo de conmutación es esencialmente infinita: no fluye corriente a través del conmutador y, por tanto, está en corte *(off)*.

Ejemplo típico de circuito lógico es el circuito negativo "no-o". Este tiene 3 entradas, 3 dispositivos de conmutación (cada uno de ellos asociado a una entrada) y una sola salida, que está situada entre el dispositivo de carga y los tres dispositivos de conmutación conectados en paralelo. Un circuito "no-o" se integra en un conjunto de circuitos (como es el caso de un procesador) conectando su salida a una entrada de uno o más circuitos similares y conectando cada entrada a la salida de uno o más circuitos distintos. Si la tensión de una cualquiera de las entradas al circuito lógico es alta (correspondiente a un 1 binario), su conmutador se cerrará y la salida será baja (correspondiente a un 0 binario). Si todas las entradas del circuito son bajas (0), todos los dispositivos de conmutación serán esencialmente circuitos abiertos y la salida será alta (1). Cabe diseñar la lógica de todo un ordenador usando sólo circuitos "no-o".

El circuito lógico disipa energía y tiene una demora de conmutación pequeña pero importante, debida al hecho de que la salida no puede responder inmediatamente a una modificación en las entradas. Tanto la cantidad de energía que se pierde como la longitud de la demora de conmutación dependen de la impedancia del dispositivo de carga. Una alta impedancia de carga da como resultado una corriente menor y, por tanto, disipará menos energía, pero una corriente menor requiere más tiempo para la apertura o cierre del conmutador. En realidad, para cualquier diseño específico de circuito, el producto de la disipación de energía y la demora de conmutación (el producto energía-demora) es una constante dentro de un rango bastante amplio y la desventaja puede combinarse con el beneficio del otro factor.

El producto energía-demora de los circuitos lógicos ha mejorado considerablemente con la aparición de los dispositivos microelectrónicos. Los dispositivos más pequeños son por sí mismos más rápidos que los grandes. Un circuito lógi-



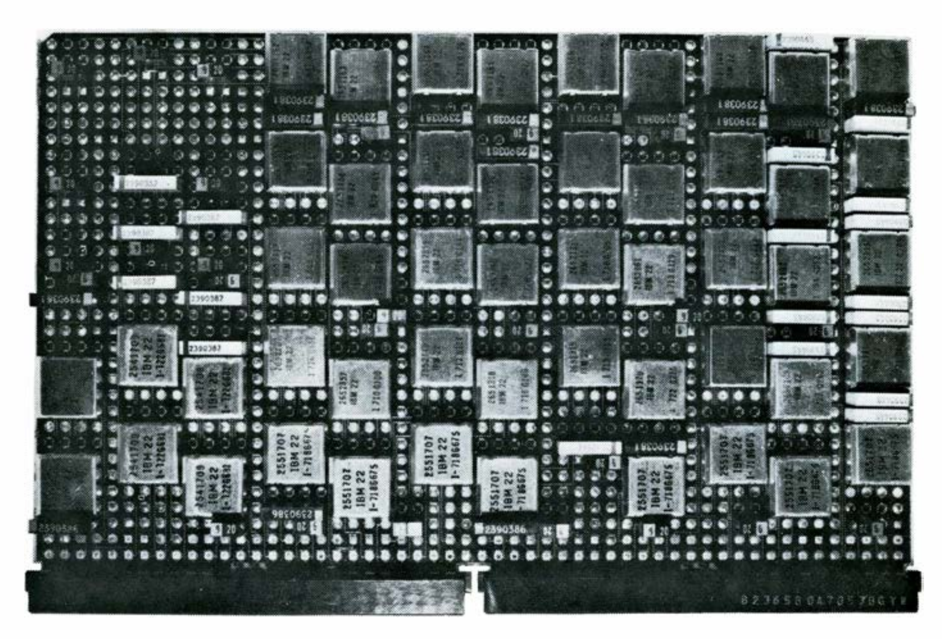
GRAN ORDENADOR CENTRAL, denominado sistema/370 Modelo 168, fabricado por International Business Machines Corporation, que es el mayor de los miembros de la familia de los ordenadores IBM/370. Está diseñado para manejar tareas a gran escala y a alta velocidad, en el campo de la ciencia o de la empresa. El juego de armarios de la parte posterior contiene la unidad central de proceso, la memoria principal y los alimentadores eléctricos. La mesa con la pantalla de tubo de rayos catódicos, el teclado y los conmutadores es la consola del operador.

co de dimensiones más pequeñas presenta una carga menor al dispositivo de conmutación y la corriente puede conmutarla con más rapidez. Es interesante indicar con exactitud cuánto ha mejorado la demora de conmutación a medida que ha progresado la tecnología de circuitos. A mediados de los años 50, el circuito lógico de un gran tubo de vacío presentaba una demora de conmutación de un microsegundo (una millonésima de segundo). Al principio de los años 60, una tarjeta de circuitos impresos, de varias decenas de milímetros de lado con transistores discretos y otros componentes montados en ella, presentaba una demora normal de conmutación de 100 nanosegundos (100.000 millonésimas de segundo). Actualmente, una pastilla corriente de circuito lógico integrado para un sistema central no llega a los pocos milímetros de lado, contiene de cinco a 10 circuitos y presenta una demora media de conmutación de menos de cinco nanosegundos. Son posibles circuitos lógicos que resulten incluso más rápidos.

Hoy día, en los circuitos lógicos semiconductores, se utilizan básicamente dos tipos de dispositivos: dispositivos bipolares y dispositivos consistentes en transistores de efecto de campo de óxido metálico semiconductor (MOSFET), (véase "Elementos del circuito microelectrónico", de James D. Meindl, en este mismo número).

Los circuitos formados por dispositivos bipolares son de fabricación más compleja y necesitan pastillas de mayores dimensiones. Los dispositivos bipolares tienen una impedancia baja y, por tanto, son rápidos. Se instalan en la unidad lógica de los ordenadores centrales y en la memoria intermedia, es decir, en aquellas zonas en que los requisitos son una alta velocidad y un buen rendimiento. Los circuitos compuestos de dispositivos MOSFET son más compactos y de fabricación más sencilla. Son relativamente lentos, pero su coste es menor, puesto que pueden empaquetarse muy densamente. Los dispositivos MOSFET se encuentran, por lo general, en la memoria principal de ordenadores y en los microprocesadores, áreas que exigen un coste bajo y en donde puede aceptarse un rendimiento menor. Se utilizan extensamente para la lógica de bajo coste en una amplia gama de aplicaciones.

La reducción de la demora de conmutación por la introducción de dispositivos microelectrónicos se refleja en un mayor rendimiento de las máquinas de proceso de datos de todos los tamaños. Por ejemplo, el tiempo de ciclo (tiempo requerido para ejecutar una operación) de un ordenador central ha disminuido de 150 nanosegundos a principios de los años 60, a 10 o 20 nanosegundos en el día de hoy, y el tiempo de ciclo de un microprocesador ha bajado de unos dos microsegundos en 1971 a medio microsegundo o menos en la actualidad.



TARJETA LOGICA DEL ORDENADOR IBM 370/168. Uno de los varios centenares de tales cuadros que componen la unidad central de proceso. Cada cuadro tiene una anchura de 11 por 18 centímetros de longitud. Lleva montados módulos de circuitos integrados *(cuadrados en plata).*

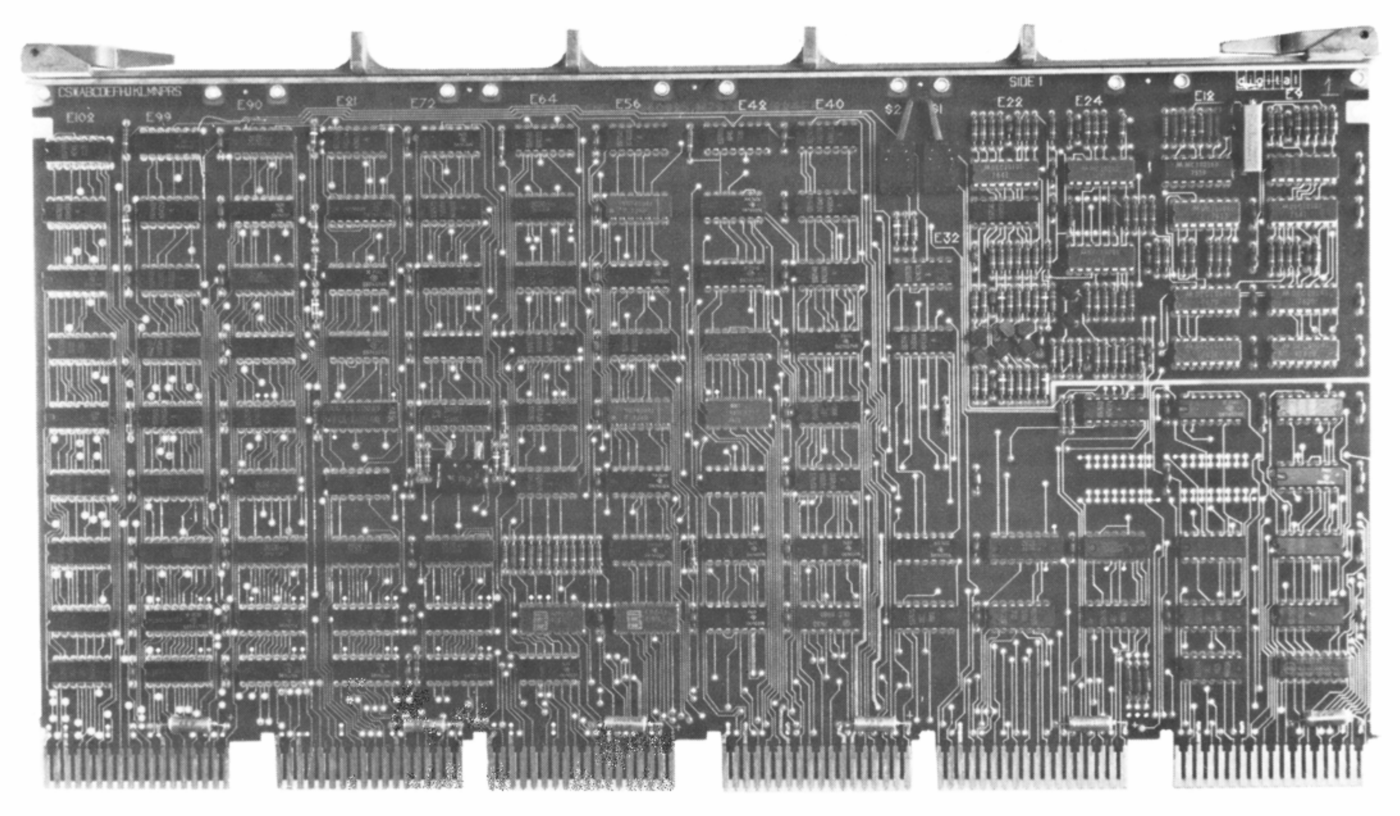
Fabricar los circuitos con elementos más pequeños no tiene sólo la ventaja de un producto más favorable en cuanto a energía-demora sino que permite, además, colocar un mayor número de circuitos en una determinada superficie de la pastilla. Aumentar el número de circuitos contenidos en una sola pastilla significa un método más eficiente de fabricar ordenadores; se necesitan menos pastillas, las conexiones entre los circuitos de una pastilla son más cortas y representan menos carga para el sistema que las interconexiones entre pastillas. Cuando los circuitos de unas pastilla han de dirigir circuitos localizados en otras pastillas, las demoras de conmutación son mayores.

Lo que limita el tamaño de los distintos elementos del circuito y de la pastilla son los "productos finales" de los procesos de fabricación. En primer lugar, para que una pastilla resulte aceptable todos los elementos que la componen deben funcionar adecuadamente. Un solo defecto en un área crítica anularía el circuito e inutilizaría toda la pastilla. Una pastilla mayor tendrá un área crítica más grande y la probabilidad de que

			IBM-370/115	IBM-370/168
UNIDAD CENTRAL DE PROCESO	NUMERO DE PASTILLAS LOGICAS		1800	20.000
	TIEMPO DE CICLO CPU		480 NANOSEGUNDOS	80 NANOSEGUNDOS
MEMORIA	INTER-MEDIA	TECNOLOGIA	—	SEMICONDUCTOR BIPOLAR
		CAPACIDAD	—	32.768 OCTETOS
		TIEMPO DE CICLO	—	80 NANOSEGUNDOS
	PRINCI-PAL	TECNOLOGIA	MOSFET	MOSFET
		CAPACIDAD	HASTA 393.216 OCTETOS	HASTA 8.368.608 OCT.
		TIEMPO DE CICLO	480 NANOSEGUNDOS	320 NANOSEGUNDOS
OTROS	TAMAÑO DE CPU		0,75 METROS LARGO 1,5 METROS DE PROFUNDIDAD 1,5 METROS DE ALTO	3,95 METROS DE LARGO 3,05 METROS DE PROFUNDIDAD 1,98 METROS DE ALTO
	PESO DE CPU		816,47 KILOGRAMOS	2313,32 KILOGRAMOS
	PRECIO DE COMPRA		175.000 DOLARES	4.500.000 DOLARES

TABLA DE INFORMACION que presenta varias estadísticas significativas sobre el menor (Modelo 115) y el mayor (Modelo 168) de los miembros de la familia IBM 370. Indica la gama hoy día disponible en cuanto a potencia de cálculo de ordenadores centrales grandes y pequeños.

MINIORDENADOR PDP-11Modelo 60, fabricado por Digital Equipment Corporation. Es uno de los miembros de tamaño medio dentro de la familia de ordenadores PDP-11. El PDP-11/60 que presentamos aquí consta de una unidad central de proceso *(cuadro de control medio contenido en el armario grande),* **con dos unidades de discos a los lados y una consola del operador** *(teclado en primer plano).* **El operador de la derecha sujeta una cápsula de discos de este ordenador. El operador de la izquierda está sometiendo a prueba el ordenador con un osciloscopio.**



CUADRO LOGICO del miniordenador PDP-11/60. Se trata de uno de los seis cuadros que posee la unidad central de proceso. El cuadro tiene una anchura de 22 por 39 centímetros de longitud. Contiene 107 circuitos integrados (dos memorias ROM) y componentes discretos.

tenga un defecto que anule el circuito es, asimismo, más elevada.

En segundo lugar, en las pastillas lógicas existe una limitación práctica en cuanto al número de circuitos lógicos aleatorios que merece la pena incluir en una sola pastilla. Una pastilla de memoria es una serie regular de celdas y circuitos de soporte. La función básica de todas las pastillas de memoria es la misma. Las conexiones de entrada-salida son sencillas. Un sistema de una sola memoria repite muchas veces un tipo específico de pastillas de memoria y el mismo tipo de pastilla puede encontrarse también en las memorias de otros sistemas de ordenador. El resultado es que se utiliza ampliamente un mismo tipo de pastillas de memoria, y el coste por pastilla es bajo.

En el caso de los circuitos lógicos, la situación es muy distinta. La lógica es mucho más aleatoria. Una determinada configuración de circuitos lógicos puede aparecer tan sólo una vez en cada máquina y puede no repetirse en esa máquina ni en ninguna otra. La situación empeora a medida que aumenta el número de circuitos de una pastilla. El resultado es que determinados tipos de pastillas lógicas aleatorias no se utilizan ampliamente y el coste por unidad es alto.

			PDP-11/03	PDP-11/70
UNIDAD CENTRAL DE PROCESO		NUMERO DE PASTILLAS LOGICAS	4	600
		TIEMPO DE CICLO CPU	3,5 NANOSEGUNDOS	300 NANOSEGUNDOS
MEMORIA	INTER-MEDIA	TECNOLOGIA	—	SEMICONDUCTOR BIPOLAR
		CAPACIDAD	—	2048 OCTETOS
		TIEMPO DE CICLO	—	240 NANOSEGUNDOS
	PRINCI-PAL	TECNOLOGIA	NUCLEO DE MOSFET	NUCLEO
		CAPACIDAD	HASTA 57.344 OCTETOS	HASTA 4.096.000 OCTETOS
		TIEMPO DE CICLO	NUCLEO: 1,15 MICROSEG. MOSFET: 750 NANOSEG.	1,26 MICROSEGUNDOS
OTROS		TAMAÑO DE CPU	482,60 MILIMETROS DE LARGO 342,90 MILIMETROS DE PROFUNDIDAD 88,90 MILIMETROS DE ALTO	533,40 MILIMETROS DE LARGO 787,40 MILIMETROS DE PROFUNDIDAD 152,40 MILIMETROS DE ALTO
		PESO DE CPU	15,87 KILOGRAMOS	226,80 KILOGRAMOS
		PRECIO DE COMPRA	2.000 DOLARES	63.000 DOLARES

TABLA DE INFORMACION que presenta estadísticas significativas sobre el menor (Modelo 03) y el mayor (Modelo 70) de los miembros de la familia de miniordenadores PDP-11. El PDP-11/70 puede considerarse en pequeño ordenador central. Tiene una memoria intermedia de alta velocidad. Este y el PDP-11/60 son los únicos miembros de la familia PDP-11 dotados de dos memorias.

Además, al aumentar los circuitos lógicos de una pastilla, se hace más difícil diseñarla. Este problema aparece no sólo al diseñar inicialmente la pastilla, sino también al diseñarla de nuevo para modificarla o corregir errores del proyecto inicial. Realizar un nuevo diseño de una pastilla lógica aleatoria puede resultar muy costoso si los circuitos de la pastilla son numerosos. Al aumentar el número de circuitos lógicos de la pastilla, crece considerablemente el número de conexiones de entrada-salida, haciendo más difícil colocar la pastilla en el ordenador. Por tanto, el aumento del número de circuitos de una pastilla no implica sólo ventajas.

El diseñador que desarrolla la red de circuitos lógicos de un ordenador puede optar entre diversos métodos alternativos. El primero de ellos es aquel en que cada circuito puede diseñarse de acuerdo con las necesidades del cliente y colocarse individualmente en una pastilla según resulte necesario, para lograr así un rendimiento óptimo. Este método de crear la lógica del ordenador en razón de las necesidades del cliente tiene todos los inconvenientes antes mencionados, pero permite un rendimiento óptimo y la más alta densidad de circuitos. También puede ser el de menor coste cuando las distintas pastillas tienen una demanda lo suficientemente alta como para amortizar el coste de su diseño.

El segundo método consiste en adaptar cada circuito al cliente y colocarlo en una pastilla con arreglo a una disposición regular, como los recuadros de un tablero de ajedrez, que incluye las vías de cableado para las conexiones entre circuitos. Una organización tal atenúa las dificultades que presenta el diseño de toda la pastilla y conserva al mismo tiempo cierto grado de versatilidad para las diferentes funciones. Resulta menos eficiente en cuanto a densidad y rendimiento, puesto que ya no es posible organizar los circuitos del modo más eficaz para cada función individual. Además, como norma general, sólo se utiliza realmente una fracción de los circuitos de la pastilla.

El tercer método es diseñar, de acuerdo con las necesidades del cliente, un número determinado de pequeñas unidades funcionales de circuitos conocidas como "macros". Estas unidades pueden ser, por ejemplo, un sumador, una unidad de lógica booleana o una pila de registros. De este modo, puede llegar a crearse una "biblioteca" de macros. Al diseñar una pastilla, se construye con una serie de macros conectadas entre sí de acuerdo con un pequeño factor de lógica adaptada a las necesidades del cliente.

El cuarto método se relaciona con el hecho de que es posible facilitar el diseño de las pastillas lógicas haciendo que su lógica se asemeje a la memoria. El modo más común es utilizar una memoria de sólo lectura. Este tipo de memoria se diferencia de las de lectura/escritura en que la información se graba en ella al fabricarla y no puede modificarse. En una memoria de estas características, la secuencia de las operaciones lógicas se almacena en forma de configuraciones de bits en una secuencia de "palabras" de bits; al acceder a estas palabras, se decodifican las configuraciones de bits y se llevan a cabo las operaciones lógicas. Para decodificar y utilizar las configuraciones de bits, se requiere sólo una cantidad relativamente pequeña de lógica. La celda de memoria de sólo lectura es menor que la celda de lectura/escritura y no es volátil. Normalmente, un solo circuito lógico equivale a una cantidad que oscila entre 10 y 20 bits de memoria de sólo lectura. La lógica de la pastilla puede modificarse mediante la simple alteración de la configuración de bits de la memoria de sólo lectura; no se requiere nuevo diseño u organización de los circuitos.

El quinto y último método para el diseño de una red de circuitos lógicos consiste en implementar las funciones lógicas mediante un microprocesador. El microprocesador es la alternativa menos eficiente en términos de rendimiento y densidad, pero es la más flexible y fácil de utilizar. La modificación de sus funciones se realiza, simplemente, programándolo de nuevo. La misma pastilla puede realizar muchas funciones diferentes, por lo que tiene una alta demanda y resulta relativamente económica. Hoy, una sola pastilla de microprocesador cuesta sólo apenas de 10 a 20 dólares.

La versatilidad y flexibilidad del mi-

croprocesador han cambiado toda la arquitectura de los sistemas de ordenador modernos. Actualmente, el proceso de la información no se lleva a cabo exclusivamente en la unidad central de proceso. Existe una tendencia a distribuir más capacidad de proceso por todo el sistema de ordenador, incluyendo varias áreas con procesadores locales para ocuparse de operaciones en dichas zonas.

Por ejemplo, una vía de entrada-salida puede disponer de un controlador para regular el flujo de información que lo atraviese. En ocasiones, el controlador puede aceptar mandatos de la unidad central de proceso y emitir señales a ésta con objeto de coordinar sus operaciones con las del resto del sistema; en otros momentos, el controlador puede operar independientemente de la unidad central de proceso. Otro ejemplo de procesador local es lo que normalmente se denomina un terminal inteligente. En un terminal, un operador conecta con el ordenador y comparte el tiempo de éste con otros operadores. Originalmente, los terminales eran "mudos": se limitaban a enviar entradas del operador a la unidad central de proceso del ordenador y representar salidas en una pantalla. Actualmente, es cada vez mayor el número de terminales "inteligentes": son capaces de realizar por sí mismos procesos preliminares sencillos sobre la entrada del operador antes de comunicarse con la unidad central de proceso del operador.

Hoy día, en muchos sistemas de ordenador se conectan entre sí un número de procesadores para formar una red de proceso distribuido. Muy comúnmente, la red consta de una serie de miniordenadores, pero también pueden incorporarse a ella ordenadores de unidad principal y microordenadores. Las vías de entrada/salida y el equipo físico para transmisión de datos se consideran también parte activa de la red, siempre que sean capaces de procesar información. La tarea se fracciona entre los elementos de la red. Cada elemento trabaja independientemente durante un período de tiempo, comunicándose con otros elementos cuando sea necesario.

El proceso distribuido ofrece una serie de ventajas. En primer lugar, puesto que muchos elementos del ordenador pueden estar trabajando en distintas partes de la misma tarea, el trabajo se realiza con mayor rapidez. Segundo, si un elemento de la red presenta un fallo, se puede desplazar su carga de trabajo a otro o compartirla entre varios elementos, de forma que el conjunto de la red sea relativamente inmune a los fallos. Tercero, la

red puede ser suficientemente pequeña y hallarse en un solo laboratorio o edificio, o puede extenderse en un área muy amplia, como, por ejemplo, las diversas sucursales de un banco. La facilidad o dificultad con que cada elemento pueda comunicarse con otro afectará al grado de manipulación efectuada sobre los datos antes de su transmisión por la red. Un importante obstáculo a la hora de diseñar un proceso distribuido que resulte eficiente es la dificultad que ofrece la confección del soporte de programación del sistema, que ha de capacitar a los diversos elementos de la red para que operen e interactúen de modo eficaz.

Los sistemas de proceso distribuido pueden organizarse de diversos modos. Un gran sistema de proceso distribuido puede organizarse en una estructura jerárquica. En la cumbre de esa jerarquía aparece un solo ordenador central o primario que se comunica con los procesadores de la red situados en un nivel secundario. Estos, a su vez, pueden comunicarse con otros procesadores de nivel terciario, y así sucesivamente. En una jerarquía pura, los procesadores de cualquier nivel dado no pueden comunicarse directamente entre sí. Es necesario dirigir las comunicaciones a través del nivel inmediatamente superior.

Alternativamente, un sistema de proceso distribuido puede organizarse según un principio de igualdad. Todos los ordenadores están en un mismo nivel y pueden comunicarse entre sí en pie de igualdad. Sin embargo, excepto en el caso de redes muy pequeñas, es muy raro que cada elemento de la red pueda comunicarse directamente con todos y cada uno de los demás elementos. En lugar de esto, la estructura jerárquica y la de igualdad suelen combinarse en un sistema híbrido en que los procesadores de un nivel específico pueden comunicarse entre sí y con los del nivel inmediatamente superior.

La microelectrónica y el proceso de datos están indisolublemente ligados. El equipo físico de las máquinas de proceso de datos, de las mayores a las más pequeñas, se crea a partir de dispositivos microelectrónicos. Los avances logrados en los dispositivos microelectrónicos han aumentado el rendimiento y disminuido el coste de la maquinaria del proceso de datos. Estos avances continuarán en el futuro previsible, aunque existen signos de que, a medida que madure la industria, disminuirá el ritmo de avance. Y a medida que el cálculo por ordenador se haga más rápido y económico, el proceso de datos seguirá penetrando con mayor profundidad en las actividades diarias.

Microelectrónica e instrumentación y control

La microeléctronica hace posible que los instrumentos de medida, además de realizar mediciones, las analicen. Ha permitido también un control totalmente automático de los procesos y de la maquinaria industrial

Bernard M. Oliver

La ciencia y la tecnología dependen por igual de la capacidad para medir la enorme variedad de fenómenos que el hombre no puede determinar, por sí solo, con precisión y que, en muchos casos, ni siquiera puede detectar. Sin instrumentos que efectúen tales mediciones, no cabe pensar en análisis y previsiones eficientes; asimismo, sin dispositivos correspondientes con las lecturas de los instrumentos no puede haber control automático. Estamos asistiendo a la aparición de una nueva generación de poderosos instrumentos de medida y dispositivos de control. El elevado rendimiento de sus componentes se debe principalmente a la incorporación, como parte integral de cada instrumento o aparato de control, de un microprocesador y de algunas memorias digitales.

El principio que subyace a todo sistema de control es el del bucle o lazo de realimentación. Un ejemplo clásico lo constituye el termostato doméstico; en éste, el bucle consta del horno, el aire ambiente en la casa, un sensor de temperatura del aire y un controlador del horno. Cuando la temperatura del aire se eleva por encima de, o desciende por debajo de, puntos predeterminados, el sensor envía la información hacia el controlador, el cual, a su vez, conecta o desconecta el horno. El sensor es, de hecho, un instrumento de medida. Sin miedo a errar, puede afirmarse que la mayoría de los instrumentos de medida forman parte de un bucle de realimentación, ya que las mediciones que realizan sirven de guía para la acción que modifica las cantidades medidas. Hay muchas mediciones que cuestan llevarse a cabo o integrarse en un lazo de control de realimentación. Y ese vacío es el que han venido a llenar los nuevos instrumentos de medición basados en dispositivos electrónicos, como el microprocesador.

Para apreciar el cambio profundo que ha experimentado la tecnología de la medición en las últimas cuatro décadas, examinaremos lo que era un laboratorio de desarrollo de comunicaciones importante, hacia mediados de los años 30. Se está sometiendo a prueba un nuevo repetidor, o amplificador, para líneas telefónicas con red radioemisora. Se miden sus características de amplificación en toda la gama de frecuencias de música y de voz.

Un oscilador sintonizable genera la señal de prueba; la salida del repetidor se mide mediante un par termoeléctrico. La potencia de salida del oscilador varía con la frecuencia, y debe reajustarse manualmente para cada nueva frecuencia hasta un valor deseado, según se lee en el instrumento termoeléctrico. Después de cada reajuste, se alimenta la salida del oscilador hasta el repetidor a través de un atenuador calibrado, y la salida del repetidor se lee en el mismo instrumento termoeléctrico. A continuación, se ajusta el atenuador, reduciendo la entrada desde la unidad osciladora hasta que la pérdida del atenuador anula la ganancia del repetidor y la lectura del instrumento termoeléctrico es la misma que antes. La lectura del atenuador equivale a la amplificación que el repetidor proporciona a esa frecuencia específica.

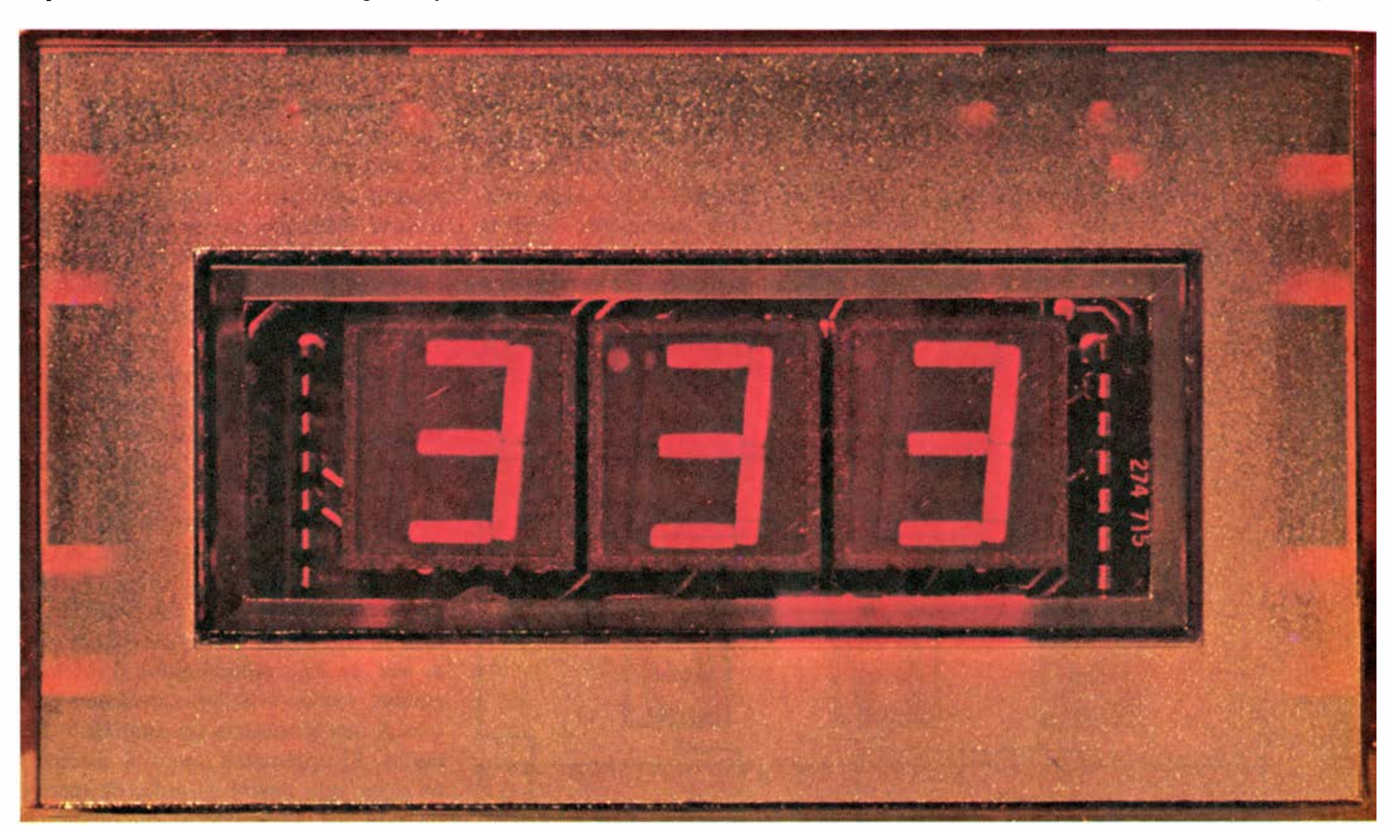
La prueba continúa. Cada nuevo ajuste de la frecuencia de oscilación debe leerse en un gráfico calibrado. Este ajuste y la lectura correspondiente del atenuador se anotan en un libro de apuntes. Después de la comida, los resultados se representarán gráficamente y se examinarán para ver si hay problemas con la respuesta frecuencial: Si los resultados son correctos, el siguiente día se dedicará a una prueba adicional del repetidor, midiendo esta vez su distorsión armónica con relación a la frecuencia.

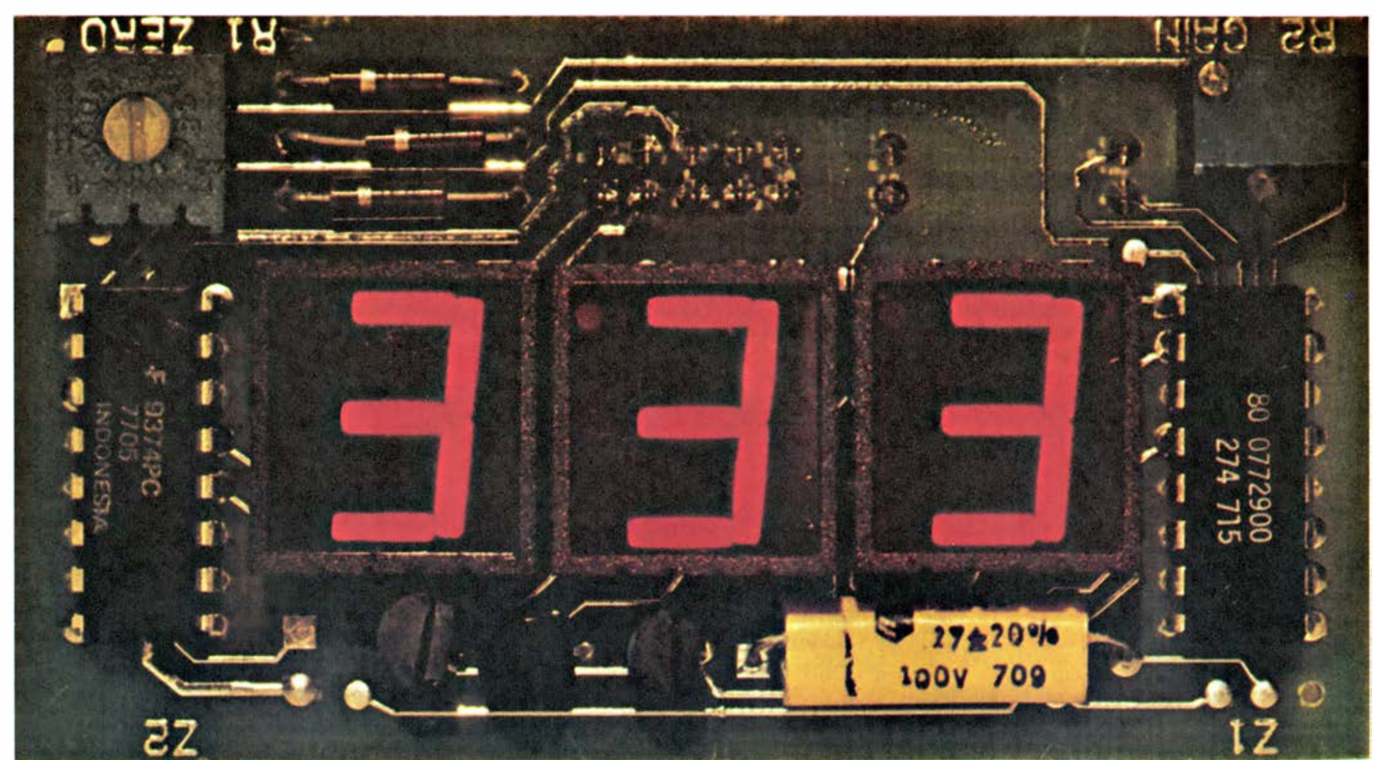
¿Cómo someteríamos a prueba hoy el mismo repetidor? Quitémosle el polvo al venerable aparato; conectemos, por cables, su entrada y su salida a un comprobador automático. Las condiciones de ensayo se imprimen en una consola y se oprime un pulsador de "MARCHA". A los 10 segundos, un terminal video refleja un gráfico que representa la ganancia del repetidor en todo el intervalo de frecuencias. La pulsación de otro botón cambia la imagen a una representación semejante de la distorsión con relación a la frecuencia. No se han girado diales, ni se han leído indicadores, pero se han resumido dos días de trabajo de toma de datos en unos cuantos segundos. Han bastado cuatro décadas para que la instrumentación progresara desde el vagón Conestoga hasta el avión a reacción. Vale la pena detenerse en el análisis de algunos avances destacados en el desarrollo de la instrumentación, a partir del año 1930 y hasta nuestros días. Sin ellos, el microprocesador no podría hoy desempeñar un papel eficaz en el campo de la instrumentación. El primer paso importante fue la estabilización de los circuitos analógicos en los instrumentos. Mediante la aplicación de circuitos automáticos de control de ganancia a los osciladores y de realimentación negativa a los amplificadores, fue posible construir generadores de señal y voltímetros con una salida y una sensibilidad constantes en todo un amplio rango de frecuencias. No hubo que recurrir a ningún método de sustitución ni ajustar atenuadores para cada lectura; en vez de ello podía ajustarse un dial para cada nueva frecuencia y leerse directamente las variaciones registradas en la amplificación.

La tecnología digital, que encontró su medio de desarrollo en los ordenadores,

se aplicó primero a la instrumentación en el aparato conocido como contador digital de frecuencias. Los primeros métodos de medición de frecuencias empleaban mucho tiempo; en cambio, un contador digital de frecuencias muestra la respuesta en una fracción de segundo y en una forma numérica fácilmente legible. De la noche a la mañana, la medición de la frecuencia se convirtió en la forma más rápida y precisa de todas las mediciones eléctricas. A ésta siguió una pronta demanda de lecturas digitales en otros instrumentos; e inmediatamente, la lectura analógica –el índice– de los voltímetros tradicionales se vería sustituida por la indicación digital.

La sustitución de las lecturas analógicas por las digitales comportó en general un número menor de errores de lectura debidos al factor humano. Y lo que es





INDICADOR DIGITAL, de rara simplicidad, que presenta instantáneamente y en pantalla numérica el valor de cualquier señal de entrada analógica que reciba, en forma de tensión o de corriente, por ejemplo, la señal de un transductor. Las generaciones anteriores de indicadores digitales incorporaban muchos más componentes, del orden de diez veces los quince que figuran en este indicador de panel AD2026, fabricado por Analog Devices Inc. de Norwood, Massachusetts. La mayoría de los componentes excitan los fotodiodos de la pantalla *(arriba)*. Hay otros dos que son los controles de entrada y un tercero que es un condensador de integración. Unos 120 componentes de los indicadores antiguos están sustituidos aquí por una pastilla de circuito integrado de gran capacidad *(abajo)*, del tipo de lógica integrada de inyección (I^2L).

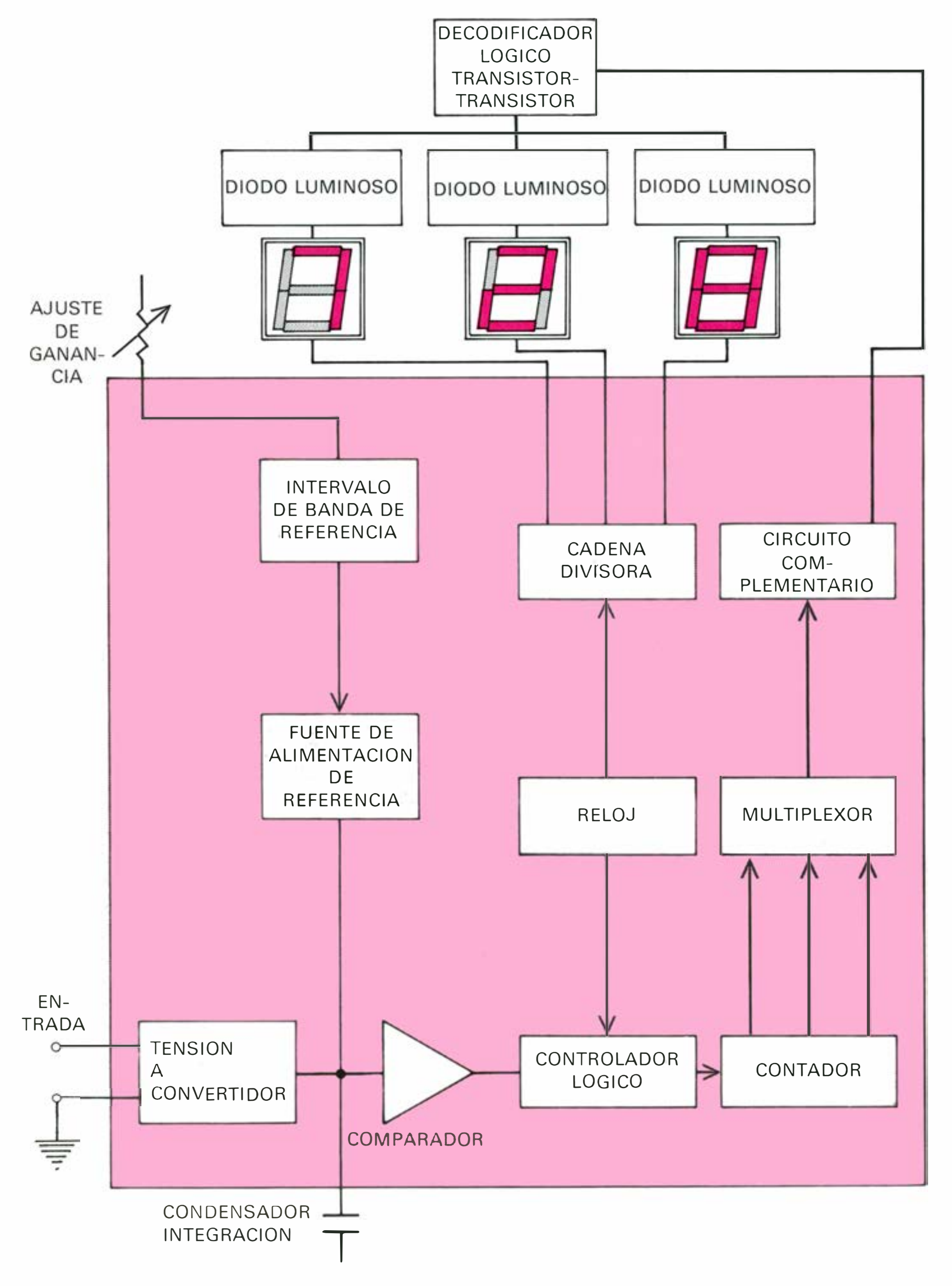
más importante, condujo a la fabricación de instrumentos cuya lectura, rápida, sin fatiga y totalmente exacta, corría a cargo del ordenador. Los ordenadores funcionaban ya para registrar los resultados de ensayos caros, que se realizaban de uno en uno; verbigracia, una explosión nuclear o el lanzamiento de un cohete. Mediante la ayuda de aparatos que convertían señales analógicas en digitales, los computadores almacenaban la salida simultánea de centenares de instrumentos, permitiendo así que los analizadores revisaran a voluntad las mediciones efectuadas. Estos primeros servicios computarizados fueron los precursores de los sistemas de medición automáticos que existen hoy.

La aparición de los instrumentos de señal de salida digital hizo que los sistemas automáticos de medición dieran un decidido paso adelante. La siguiente etapa sería el desarrollo de los instrumentos de entrada digital; es decir, de los instrumentos que se ajustaban digitalmente y que, por tanto, podían controlarse por ordenador. El contador de frecuencias quedó nivelado por el sintetizador de frecuencias y el voltímetro digital por la fuente de alimentación digital ajustable. Las fuentes de señal pasaron a depender gradualmente, una por una, del control del ordenador; todas ellas, conjuntamente con los instrumentos legibles por ordenador, constituyeron la base de una amplia variedad de sistemas automáticos de medición.

Tales sistemas sirvieron para acelerar la velocidad de las mediciones electrónicas en uno o dos órdenes de magnitud y mejoraron su precisión en una forma análoga. De manera particular, se reconoce su ventaja en la prueba de los productos industrializados, en la que se repiten una ´y otra vez ensayos rutinarios, complejos a menudo, y en donde no es frecuente la necesidad de una nueva programación.

Los sistemas de ensayo automático poseen muchas ventajas, aparte de relevar a los operarios de las tareas monótonas. El sistema automático puede compensar sus errores sistemáticos de medición, midiendo en relación a una referencia normalizada y guardando en su memoria la diferencia entre estas mediciones y los valores conocidos. Tales diferencias, que representan errores, pueden restarse cuando se mide un aparato en la línea de fabricación. Así se reduce el nivel del error en la medición en unas diez o cien veces.

Para obtener una nueva lectura, un sistema automático de medida no necesita más tiempo que el requerido para estabilizar el aparato ensayado y los instrumentos asociados. Este tiempo puede ser de sólo milisegundos o microsegundos, mucho más corto, evidentemente, que el que emplearía un analizador humano. El ordenador opera el sistema a un ritmo que supera en miles de veces el seguido por el hombre, y, en consecuencia, obtiene más datos por hora del equipo de ensayo.



CIRCUITO DE INDICADOR DIGITAL, que comprende los elementos de pastilla I²L encerrados dentro de la caja coloreada y 14 componentes adicionales. La mayoría de éstos se hallan relacionados con la indicación numérica mostrada en relieve, en la parte superior de la figura.

El sistema automático posee una capacidad específica en el proceso de datos. De los millares de mediciones que se efectúan cada minuto, el operador humano necesita sólo las que presentan alguna anomalía; mediciones éstas que pueden mostrarse completamente junto con la información del error. Los datos de ensayo obtenidos en diferentes fabricaciones pueden almacenarse para una ulterior referencia, revisarse ítem por ítem y analizarse estadísticamente para determinar si los procesos de fabricación se desarrollan bajo control.

Los sistemas de ensayo automático no falsean los datos, ni se equivocan al registrarlos, ni se cansan, ni omiten pruebas, ni caen en ninguna de las docenas de deficiencias propias de la condición humana. Cualquiera que sea la prueba que el programa determine, se realizará independientemente de la hora o del día de la semana; ni valen presiones de la oficina central para terminar y enviar el género a tiempo que puedan comprometer la labor inspectora del ordenador.

¿Qué ocurre en la etapa siguiente del desarrollo, cuando cada instrumento in-

corpora su propio ordenador bajo la forma de un microprocesador? Se abre un amplio abanico de nuevas posibilidades y de clases de rendimiento. Aparece una instrumentación "inteligente". Vamos a examinar algunos ejemplos típicos, que nos han de ayudar en la comprensión de lo que son capaces de hacer estos instrumentos inteligentes.

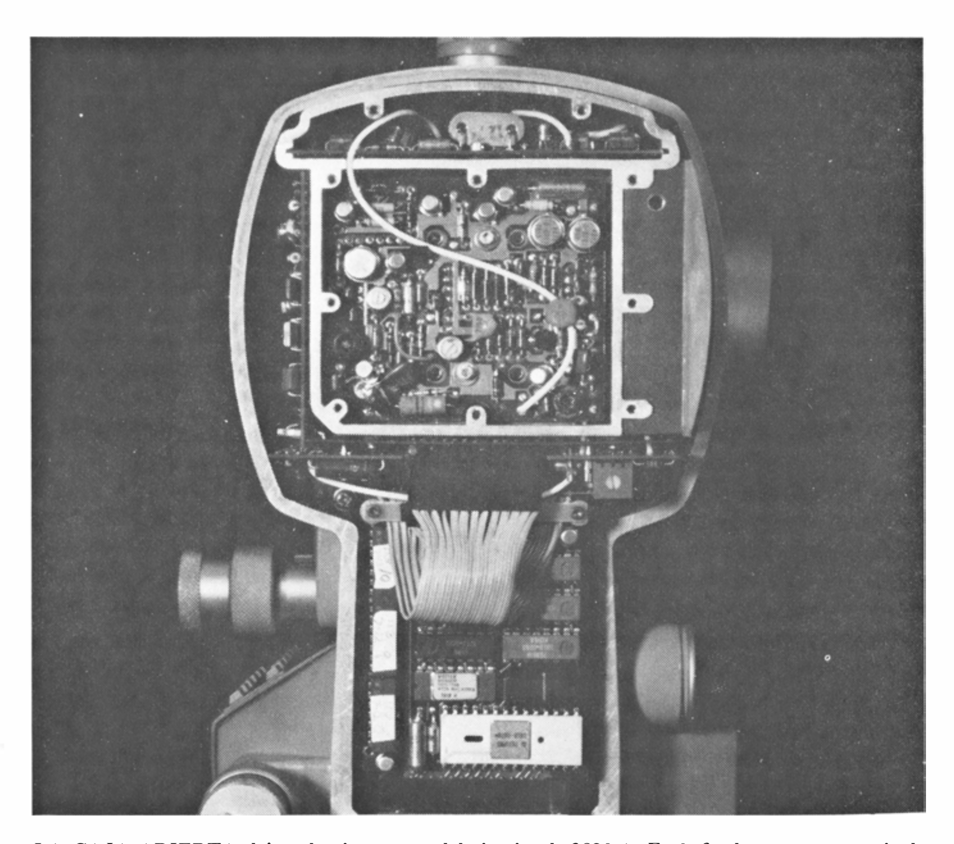
Comparemos un analizador de espectro tradicional con otro inteligente. Un analizador de espectro mide la potencia de una señal en función de la frecuencia a lo largo de una gama específica del espectro de frecuencias. Se trata de un receptor de banda estrecha que barre repetidamente una parte seleccionada del espectro y que presenta visualmente todas las señales que recibe, en forma de picos, cuya altura es proporcional a la potencia de la señal. Sin embargo, si la velocidad de barrido es muy alta, el receptor no puede responder exhaustivamente a las señales que detecta; en el extremo opuesto, si la velocidad de barrido resulta excesivamente baja, la indicación puede ser intermitente y oscilar, con lo cual pueden perderse cambios ocurridos en la señal.

Para obtener la más alta velocidad de repetición posible, sin sacrificar ni los detalles ni la precisión en la pantalla, deben optimarse tres factores que se hallan interrelacionados: selección de la anchura del espectro, resolución dentro del espectro y velocidad de barrido. Antes de introducirse la microelectrónica, la tarea de la optimación se dejaba en manos del operador; sólo un hombre cuidadoso y entendido podía obtener resultados coherentes, exactos y completos.

En los analizadores de hoy en día, el microprocesador ajusta automáticamente la resolución espectral para mostrar el mayor detalle posible que permite la pantalla; luego, selecciona la velocidad de barrido más alta que permite todavía una respuesta completa y exacta para cada señal detectada. Aun cuando la velocidad de barrido sea muy baja, la indicación no oscilará; se almacenará en una memoria digital y se presentará al operador 60 veces por segundo. El microprocesador no sólo presenta los picos de la señal sino también otros datos pertinentes: tiempo de barrido, amplitud de banda del receptor, intervalo de medida del espectro, intensidad de la señal y frecuencias mínima, máxima y central de la señal. El control fotográfico de la pantalla permite registrar todas las secuencias en cualquier momento. En vez de luchar contra las deficiencias del analizador de espectro clásico, el operador se encuentra ahora ayudado eficaz-

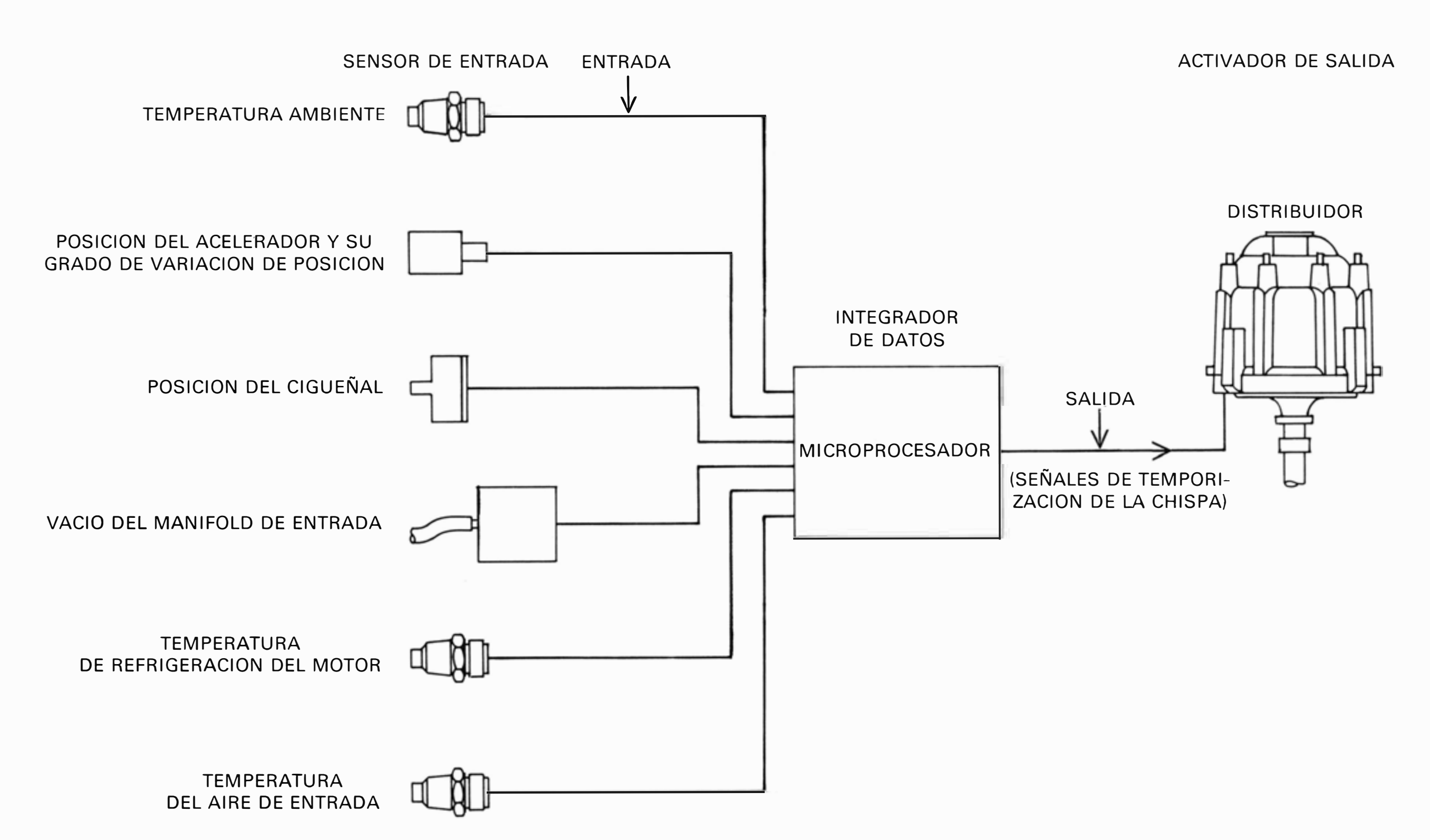


INSTRUMENTO DE EXPLORACION "INTELIGENTE", que combina las mediciones angulares de un teodolito con las mediciones a distancia obtenidas por procesamiento microeléctronico de las ondas reflejadas de luz láser. En el instrumento, un microprocesador presenta las lecturas del azimut y del ángulo de elevación en grados, minutos y segundos de arco. Un sensor interno detecta cualesquiera errores de nivelación y corrige de forma acorde las lecturas angulares. En mediciones de distancias de hasta cinco kilómetros, el microprocesador convierte las lecturas de inclinación de la línea visual en sus componentes horizontal y vertical, mientras efectúa las correcciones correspondientes a la curvatura de la tierra. El instrumento es el Hewlett-Packard 3820 A.



LA CAJA ABIERTA deja a la vista parte del circuito de 3820 A. Enchufando en una memoria de estado sólido, el explorador puede registrar un día de observaciones de campo para transferirlas a un ordenador en la central, directamente o por teléfono. El circuito de mantenimiento de datos puede conectarse a un trazador, que convertirá los días de observaciones en un trazado normalizado.

EL CONTROLADOR de motores de automóvil incorpora un microprocesador para integrar las lecturas de seis instrumentos sensibles y determinar el tiempo óptimo de ignición para cualesquiera valores dentro de un amplio campo de velocidades y de cargas del motor. El control, que está en fase de desarrollo por la sociedad Chrysler Corporation, emplea una pastilla de microcesador fabricada por la RCA Corporation.



SIETE ENTRADAS DE SENSORES en el circuito monitor de control del motor Chrysler. Comprenden la temperatura del aire ambiente, la temperatura del aire de entrada, la temperatura de refrigeración del motor, el vacío del manifold o colector de entrada, la posición del cigüeñal, la posición del acelerador y el grado de variación de la posición del acelerador. El mejor rendimiento del motor, conjuntamente con la relación fuel-aire idónea para mantener la mezcla que haga mínima la polución, depende de la selección del instante óptimo de la ignición con respecto a las variables captadas. El microprocesador envía estas señales de tiempo al distribuidor, según se refleja claramente en el gráfico.

mente. Los controles están simplificados, y se pueden almacenar valores completos en la memoria para solicitarlos más adelante en caso de tener que repetir las mediciones.

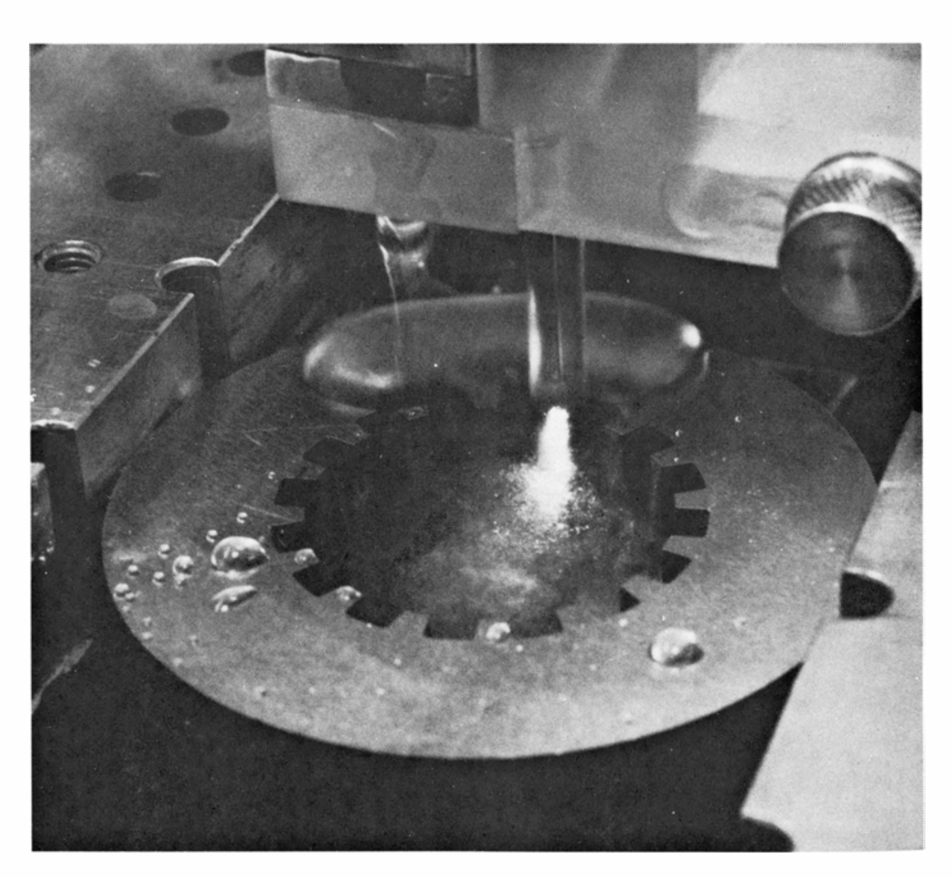
Hay un segundo ejemplo, que tiene que ver con la combinación de dos instrumentos de análisis químico: el cromatógrafo de gases y el espectrómetro de masas. Cuando una muestra de una composición desconocida pasa a través de la columna del cromatógrafo de gases, cada componente de la mezcla viaja a su velocidad característica, de forma que raramente dos componentes alcanzan el extremo de la columna al mismo tiempo. La llegada de cada componente al final de la columna se capta de múltiples modos y queda registrada cuantitativamente en forma de pico en un gráfico rectangular. Tal gráfico constituye el cromatograma. Por ejemplo, una gota de gasolina producirá un cromatograma con un pico distinto para cada hidrocarburo presente. El químico puede determinar después en qué porcentaje se encuentra cada hidrocarburo en la mezcla, mediante integración del área en cada pico, lo cual es un proceso laborioso.

Hace pocos años, los procedimientos seguidos en la cromatografía de gases recibieron un fuerte impulso al convertir la señal de salida del detector, de analógica en digital, y almacenar el cromatograma en la memoria de un computador; de este modo, el área de cada pico podía integrarse numéricamente con ayuda de un algoritmo adecuado del ordenador para deslindar el principio y el final de cada pico. Hoy en día, con los microprocesadores, cada cromatógrafo puede convertirse en un instrumento cuantitativo autocontenido (self-contained), dotado de calibración propia y de rutinas de autocomprobación internas y con su propio integrador incorporado.

En el espectrógrafo de masas, se introduce una muestra de composición desconocida en un recinto en que se ha hecho el vacío y se ioniza mediante un haz de electrones. Los iones de masa distinta se separan, de acuerdo con su relación de masa a carga eléctrica; el resultado obtenido se registra en series de picos aislados, análogos a los picos del cromatograma. Estos espectros de múltiples picos representan un conjunto de "firmas", que son características exclusivas de los distintos elementos y componentes. De este modo, se ha llegado a reunir verdaderas bibliotecas de firmas para colaborar en la identificación de muestras de composición desconocida. El defecto de este sistema reside en el solapamiento del espectro ante una mezcla de componentes, con lo cual su identificación



MAQUINA HERRAMIENTA NUMERICAMENTE CONTROLADA *(primer plano)*, **programable por ordenador. Al fondo, se encuentra el teclado y el componente de memoria. Se introducen los programas para varias tareas y se solicitan conforme se van necesitando. La herramienta opera en la Wilsey Tool Company, Inc., de Quakertown, y realiza una mecanización por descarga eléctrica; el electrodo es un hilo móvil. Está fabricada por la Andrew Engineeringof Hopkins, Minnesota.**



UNA TAREA TIPICA para un hilo móvil EDM es el corte de dientes de engranaje; en este ejemplo, los dientes están en la cara interior de la rueda de engranajes. En muchas clases de corte interior, el hilo móvil debe primero roscarse manualmente a través de un orificio de puesta en marcha. Luego, el EDM no precisa ya de supervisión. Cuando se programa una serie lenta y compleja de operaciones de corte, el sistema puede funcionar, sin mayor atención, durante 60 horas.

resulta muy difícil y carente de seguridad.

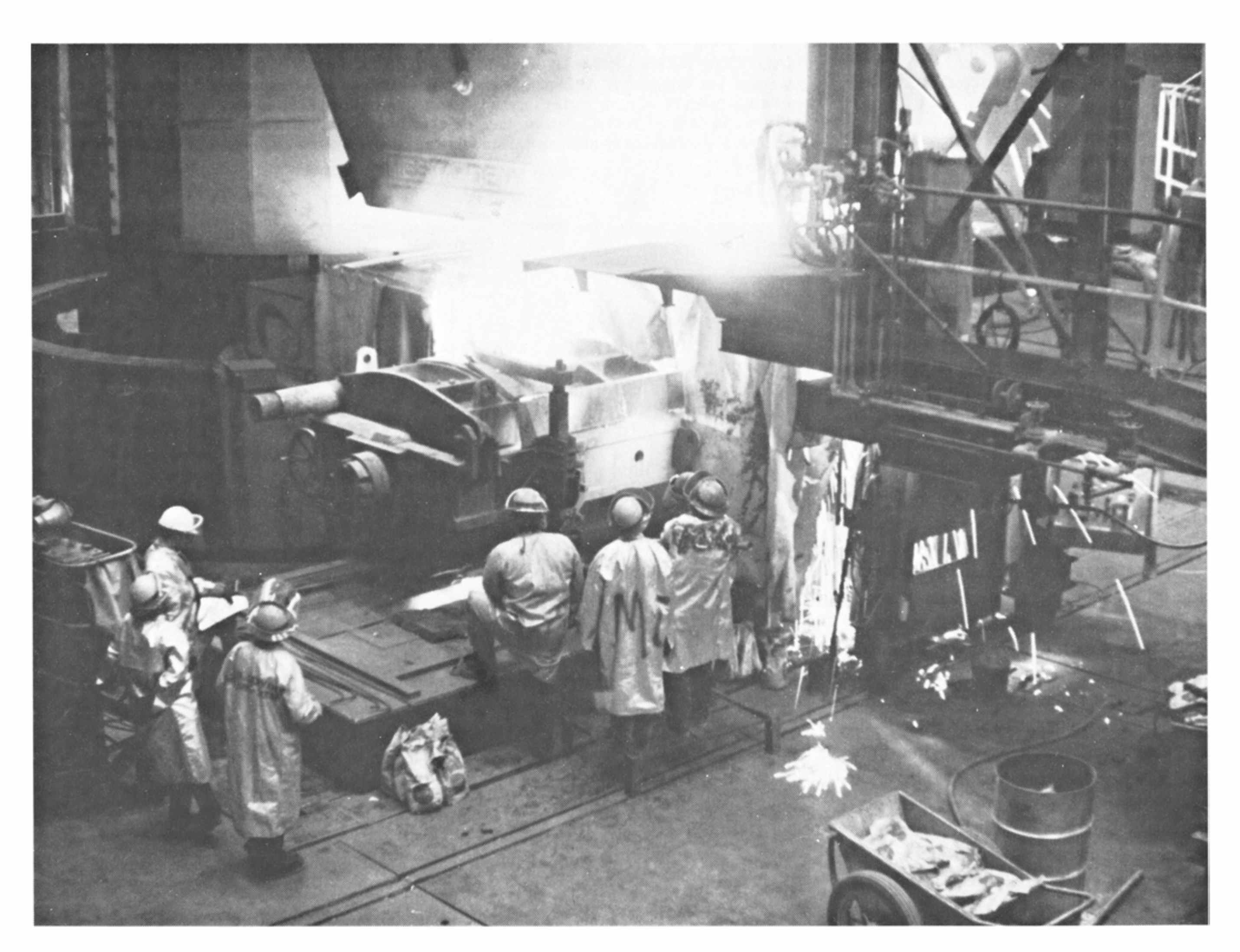
En contraposición, cada pico del cromatograma representa casi siempre una sustancia única; y ello explica la conveniencia de que se acoplen en tándem ambos instrumentos. El cromatógrafo podría separar la mezcla desconocida en sus constituyentes y el espectrómetro de masas podría identificar entonces cada producto constituyente.

La combinación anterior constituye una poderosa herramienta analítica. También es un sistema complicado que requiere una corrección y recalibración diarias para compensar la deriva resultante de la contaminación por muestras pasadas. Cuando la combinación de estos instrumentos se coloca bajo el control del microprocesador, no sólo se reduce, de horas a minutos, la tarea de recalibración y se simplifica grandemente el registro de la deriva, sino que también se puede analizar automáticamente un elevado número de muestras.

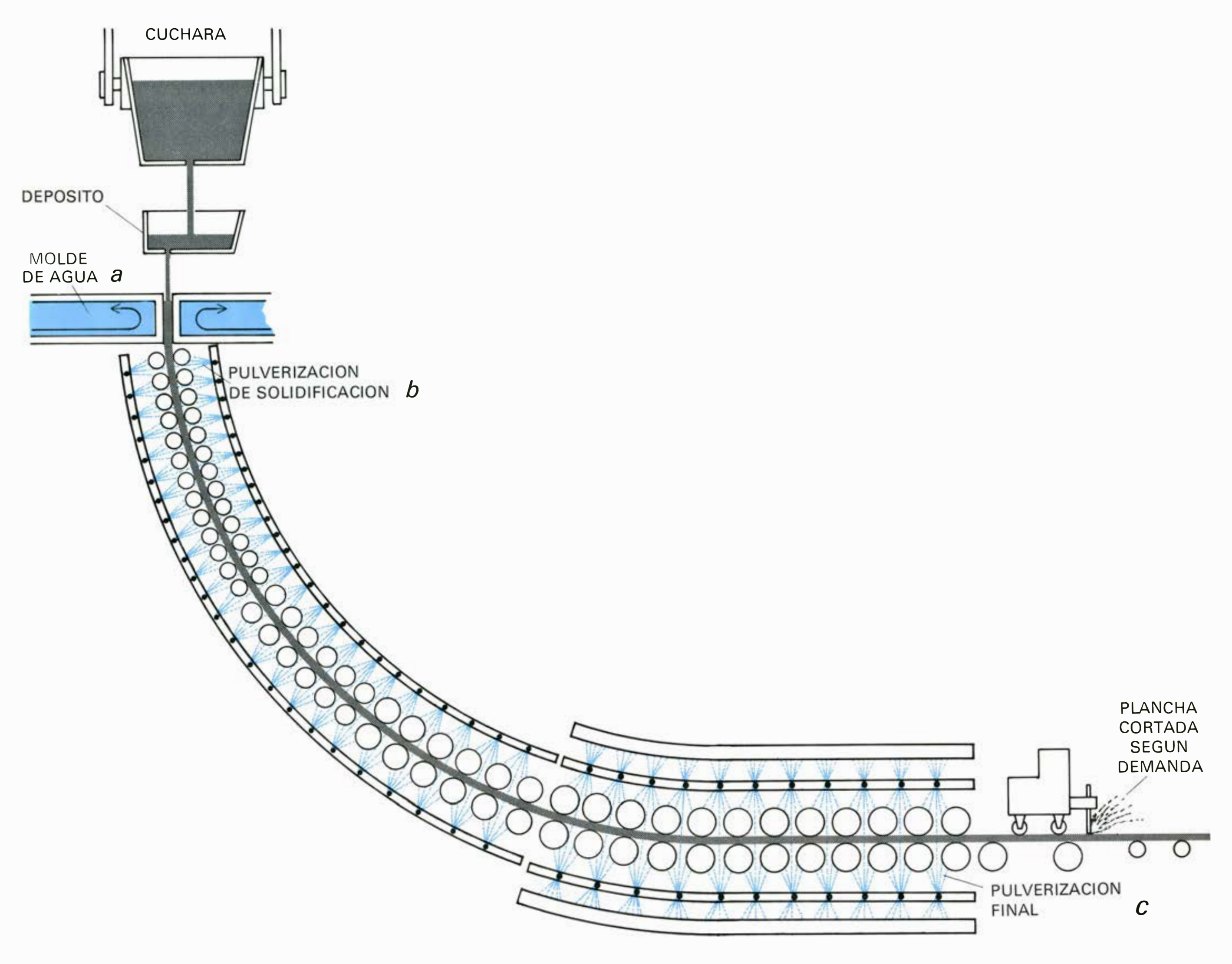
El espectrómetro de masas puede ordenarse a generar espectros para cada pico en el cromatograma o analizar solamente los espectros que satisfacen criterios predeterminados. La máquina en tándem puede trabajar toda la noche, produciendo cromatograma tras cromatograma y comprobando sus espectros de masas con las firmas almacenadas en su memoria. Si se le configura así, podrá hasta imprimir y almacenar en su memoria los "10 mejores" que resulten en cada comparación para que el operador los revise por la mañana. Nos hallamos ante una química analítica casi completamente automatizada de una clase muy sofisticada, que puede ajustarse para detectar trazas de cualquier sustancia, desde narcóticos en la orina hasta contaminantes en el aceite de limón. Todo lo que se requiere para tal tipo de automatización es el concurso de microprocesadores y memorias adecuadas.

Durante décadas se profetizó la llegada inminente de la fábrica automática, pero ésta nunca llegó. La razón fundamental estriba en que la automatización implica mucho más que el mero desarrollo de máquinas secuenciales. Necesita dispositivos de realimentación que capten las anomalías en el sistema, las analicen y tomen las oportunas medidas de corrección. En la película *Tiempos modernos,* Charlie Chaplin es alimentado, mientras está en el trabajo, por un robot que le introduce la comida en todas partes menos en la boca. Charlie era demasiado bajo para la máquina, y no podía adaptarse a ella. En un gran proceso de fabricación, hay miles de cosas que pueden errar el camino y convertir un sistema incorregible en una especie de baile.

Pasará todavía mucho tiempo antes de que la mayoría de los procesos de producción puedan automatizarse enteramente. No obstante, como las máquinas van ganando en destreza, cabe esperar que compitan con éxito dentro de una



LA FUSION CONTINUA de planchas de acero necesita una rutina, automática y compleja, en la refrigeración por agua. Esta máquina es el moldeador de hilo continuo y radio constante en el complejo Burns Harbor, Indiana, de la Bethlehem Steel Corporation; la plancha mide más de 1,8 m de ancho y puede tener un espesor de 0,2 a 0,3 m. La temperatura del agua, la presión de la válvula y la variación de caudal, a lo largo de las sucesivas etapas, se mantienen dentro de límites predeterminados mediante el sistema SPEC 200 de Foxboro Company.



LA PLANCHA DE ACERO se enfría primero en el molde *(a)*, **donde el acero fundido empieza a solidificarse. Luego, se enfría la plancha descendente, mediante agua pulverizada a ambos lados** *(b)*. **Asimismo, el armazón principal y los rodillos de la plancha plana** *(c)* **se refrigeran con agua. Los sensores de temperatura, de presión y de caudal suministran los datos de entrada para los bucles de realimentación del SPEC 200.**

gama cada vez más amplia de problemas. Buen ejemplo lo constituye la evolución de los sistemas de señalización del tráfico. Las primeras señales automáticas de tráfico eran puramente secuenciales; tenían una temporización fija y operaban con independencia una de otra. El tráfico potenció su fluidez, hace muchos años, mediante la sincronización del sistema y cambiando su temporización para acomodar las variaciones de intensidad que se producen en las horas punta. Aun así, cuando el tráfico se volvía anormalmente denso, había que acudir a la regulación "manual" de los agentes urbanos.

Hoy en día, las señales de tráfico pueden disponerse para captar y contar el número de coches en la vía principal de tránsito, e incluso la acumulación de vehículos en las calles de acceso. Pueden programarse para ajustar su propia temporización al mantenimiento en un punto óptimo del flujo del tráfico; el control adaptativo de esta clase es una tarea trivial para el microprocesador. Por otra parte, una vez se ha incorporado el microprocesador en la señal de tráfico, simplifica la comunicación con un controlador central. Cuando las señales de tráfico "inteligentes" pasan a ser ubicuas y enlazan con un control central, el programador de tráfico sito en la intersección pierde su razón de ser.

En las máquinas herramientas puede verse una evolución muy afín. Tales aparatos (por ejemplo, los tornos de torreta giratoria y las máquinas de extrusión) han llevado a cabo operaciones secuenciales complejas por mucho tiempo. El "programa" formaba parte del soporte físico: levas y engranajes mandaban secuencialmente la máquina a través de una serie de pasos para fabricar una pieza particular. Sin embargo, la variedad de secuencias estaba limitada y el coste de los nuevos programas de soporte físico era elevado.

En los primeros años de la década de los 60, aparecieron las máquinas herramientas de control numérico: máquinas acopladas a lo que se llamó un miniordenador. Tales máquinas podían tallar, perforar, taladrar y aterrajar orificios en cualquier ángulo en una pieza, sin apenas importar su forma. La herramienta adecuada venía seleccionada automáticamente mediante un programador de cinta perforada que leía un controlador-ordenador electrónico. El mismo programa situaba la herramienta en las tres coordenadas y señalaba la velocidad del eje de la herramienta y la profundidad y velocidad de corte. Para revisar la forma final de la pieza, o para hacer una pieza totalmente distinta, bastaba con reformar la cinta perforada usada o fabricar otra nueva.

Las máquinas de control numérico de esta clase llenaron una necesidad importante en la fabricación de volumen intermedio, donde la salida de las piezas

fabricadas es substancial, pero no lo suficientemente alta como para garantizar una línea de producción masiva (en donde cada serie de máquinas lleva a cabo una sola operación). Sin embargo, las máquinas acopladas a un ordenador resultaban bastante caras, radicando una parte nada despreciable del coste en el controlador-ordenador. Hoy día, los controladores basados en el microprocesador, de tamaño no mayor que las calculadoras de mesa, están ocupando el lugar de los grandes controladores de la primera generación; y asimismo, los cassettes están sustituyendo a las bobinas de cinta perforada. El coste de los controladores microelectrónicos no agrega gran cosa al coste de la máquina herramienta.

Los controladores basados en los microprocesadores permiten también un acceso por teclado a la máquina herramienta. Cuando un futuro taller se equipe con máquinas de esta clase, cada una de ellas dotada con programas almacenados que puedan activarse a distancia mediante un teclado, el ordenador central podrá dirigir la producción automática total. La única intervención humana que se requerirá será la carga de la pieza en un extremo, el transporte de las piezas terminadas en el otro extremo y la sustitución periódica de las cuchillas gastadas.

Cabe esperar la aparición de progresos ulteriores en los controladores electrónicos para automóviles. El microprocesador hace ahora económicamente posible la extensión de la aplicación de los controles electrónicos a tales funciones, como la temporización de la ignición y la carburación. Es muy fuerte el incentivo por desarrollar tales sistemas; en efecto, no sólo favorecen la economía de combustible, sino que, además, ayudan en el cumplimiento de las normas sobre emisión de gases de motor.

La temporización correcta de la ignición es función de la velocidad del motor y del vacío en el colector (manifold) de entrada. Los dispositivos que se emplean hoy en los coches sólo pueden aproximarse a los puntos óptimos, mientras que la aplicación de los microprocesadores puede asegurar la temporización correcta para todas las velocidades y cargas. En teoría, estos aparatos deben desempeñar el papel de monitores, sin asumir la misión de suministrar cada carga de combustible y de encender cada chispa. Tal monitor compararía el rendimiento real con el ideal y efectuaría los necesarios ajustes en el sistema básico existente. Si se utilizara así, el microprocesador tendría la ventaja de manejar una tasa de datos muy exigua, mandando sólo unas pocas correcciones por segundo. Además, sería seguro en caso de fallo: si el microprocesador fallara, el motor continuaría en marcha, viendo simplemente disminuido su rendimiento al nivel de los actuales.

El propio microprocesador podría registrar continuamente el estado de funcionamiento del vehículo y señalar en una pantalla LED las averías o contratiempos que se presentaran, verbigracia: bajo caudal de combustible, calentamiento del motor y calentamiento o merma de frenos. El coste de tales sistemas reside no tanto en los componentes microelectrónicos cuanto en los sensores necesarios para suministrar los datos al microprocesador y en los accionadores convenientes para ejecutar las instrucciones del mismo.

Los sistemas de control y sus componentes sensibles asociados tienden cada vez más a parecerse al sistema nervioso humano, según se deduce. Señalemos que los microprocesadores instalados en los controladores de satélites, al permitir que la "inteligencia" se amplíe a través del sistema de control, reducen mucho la cantidad de datos que deben transmitirse, vía los enlaces de datos, para ser manejados por el ordenador central. En el sistema nervioso humano, una porción notable del control se encuentra a nivel local y buena parte del proceso de los datos de entrada se hace en los ganglios. El sistema nervioso es como una red de sensores y de microprocesadores conectados por enlaces de datos a un ordenador central. Ambos, los procesadores de satélites y el ordenador central, poseen una gran cantidad de soporte firme (firmware); es decir, de subrutinas que nos permiten realizar el 99 por ciento de lo que hacemos sin "pensar", sin acudir a los centros de control superiores. Aquí, los sistemas de control microelectrónicos se parecen también al sistema de control humano: el ordenador central debe mandar que se opere, sin determinar a las máquinas más pequeñas cómo deben hacerlo.

Los ejemplos que he citado son una muestra sumaria de los cientos de avances con que el microprocesador mejorará y simplificará la operación de las máquinas que nos sirven cada día. La tecnología de los circuitos integrados ha reducido de tal modo el coste de la computación y de los sistemas lógicos que la mayoría de los aparatos que empleamos pueden ser ahora mucho más cooperantes e inteligentes que lo fueron en el pasado. Estas décadas se reconocerán en el futuro como la era inicial de la revolución robótica.

Microelectrónica y comunicaciones

Los sistemas como la red telefónica, la radio y la televisión tienen su fundamento en el procesamiento de la señal. Por su gran capacidad y bajo coste, los dispositivos microelectrónicos son ideales para ellos

John S. Mayo

Hoy se puede disponer de un teléfono que "recuerde" los números más solicitados, que, con sólo apretar un botón, obtenga la comunicación con cualquiera de ellos. En la televisión se pueden ver acontecimientos ocurridos en cualquier rincón de Europa y del mundo entero en el mismo momento en que se producen. Los ordenadores hablan "digitalmente" entre sí a increíbles velocidades, utilizando las líneas telefónicas. Estos avances y muchos más reflejan la incidencia de la microelectrónica en el ámbito de las comunicaciones.

Hasta hace 15 años, el tubo de vacío era el componente activo que dominaba en los circuitos electrónicos, fundamentales para el funcionamiento de los sistemas de telecomunicación. Sus limitaciones quedaron en evidencia ante el rápido desarrollo de las comunicaciones por teléfono, radio y televisión. El tubo de vacío era demasiado grande, necesitaba demasiada potencia y resultaba demasiado poco fiable para cumplir las exigencias, por parte de estos sistemas, de un gran número de dispositivos de proceso de señal agrupados en complejos circuitos y que debían funcionar con una fiabilidad extrema. La alternativa vino con el transistor, que aportaba ganancia electrónica dentro de un semiconductor y era pequeño y fiable. De él se pasó, de forma natural, al circuito integrado de silicio, que supuso una revolución en la electrónica y un enorme progreso en las telecomunicaciones.

Según han demostrado los artículos precedentes, hoy se fabrican simultáneamente millones de elementos de circuito sobre una fina "oblea" de silicio. Esta suele contener varios cientos de copias de un mismo circuito; se divide en pequeñas pastillas que contienen un circuito cada una. Las pastillas se encapsulan individualmente, para dar lugar a un componente llamado circuito integrado. El diseñador de equipos de comunicación emplea estos componentes para construir sistemas complejos, generalmente colocando un cierto número de circuitos integrados sobre una tarjeta de circuito impreso que contiene el cableado necesario para interconectarlos.

Un sistema de comunicación típico recurre a distintas tecnologías de fabricación de circuitos. Por ejemplo, las señales transmitidas con alta potencia, o las de alta frecuencia, usan a menudo corrientes y tensiones que no pueden ser tratadas por circuitos integrados clásicos. Algunos circuitos precisan componentes altamente estables; frecuentemente, la mejor manera de obtenerlos es mediante técnicas de capa fina, esto es, depositando delgadas capas conductoras, aislantes y resistivas sobre substratos cerámicos. Para describir con exactitud cuál es el papel de la microelectrónica en las comunicaciones, examinaré algunas de las tecnologías más importantes de fabricación de circuitos, haciendo especial hincapié en la de los circuitos integrados de silicio, por ser la que, con gran diferencia, ejerce un mayor influjo en las telecomunicaciones modernas.

La microelectrónica en forma de circuitos integrados es un factor de importancia en telecomunicación y ello, fundamentalmente, en virtud de los efectos combinados de bajo coste, alta fiabilidad, y gran versatilidad. A medida que aumenta el número de elementos de circuito fabricados conjuntamente en una misma pastilla, el coste de una función de circuito básica disminuye considerablemente. Una función de circuito de extraordinaria importancia para los sistemas de comunicación (y también para los de informática) es la "puerta" lógica digital. Esta controla el flujo de información, dando una señal de salida sólo cuando las señales de entrada están en situaciones previamente determinadas. A partir de este elemento básico pueden construirse grandes sistemas de proceso de señal. En consecuencia, el coste de una puerta lógica tiene una considerable influencia en el coste de los equipos de comunicación; esto es: en los terminales, que actúan de interfaz entre las personas o las máquinas y los canales de comunicación; en las centrales, que establecen los caminos de comunicación; por último, en los equipos que procesan la señal, de forma que puede transmitirse por líneas y cables, por radio y mediante ondas de luz.

La tecnología digital ha progresado rápidamente desde las puertas lógicas realizadas con tubos de vacío a las transistorizadas, para pasar poco después a las puertas lógicas integradas y, por último, a los circuitos integrados, con miles de puertas cada uno. Durante esta evolución, el coste ha descendido desde unos 10 dólares para la puerta de tubo de vacío hasta alrededor de un centavo para cada una de las puertas de un gran circuito integrado. Este coste bajará próximamente a los 0,1 centavos. Con esta reducción de coste en más de mil veces, de un circuito crucial, los equipos que suponen un complicado proceso de la señal, como es el caso del teléfono con memoria, resultan viables desde el punto de vista económico.

No basta, sin embargo, con un coste bajo. Un canal de comunicación puede llegar a necesitar que miles de circuitos trabajen simultáneamente. El fallo de cualquiera de sus elementos puede romper la comunicación. En resumen, la fiabilidad resulta tan importante como el coste.

De nuevo la puerta lógica sirve de

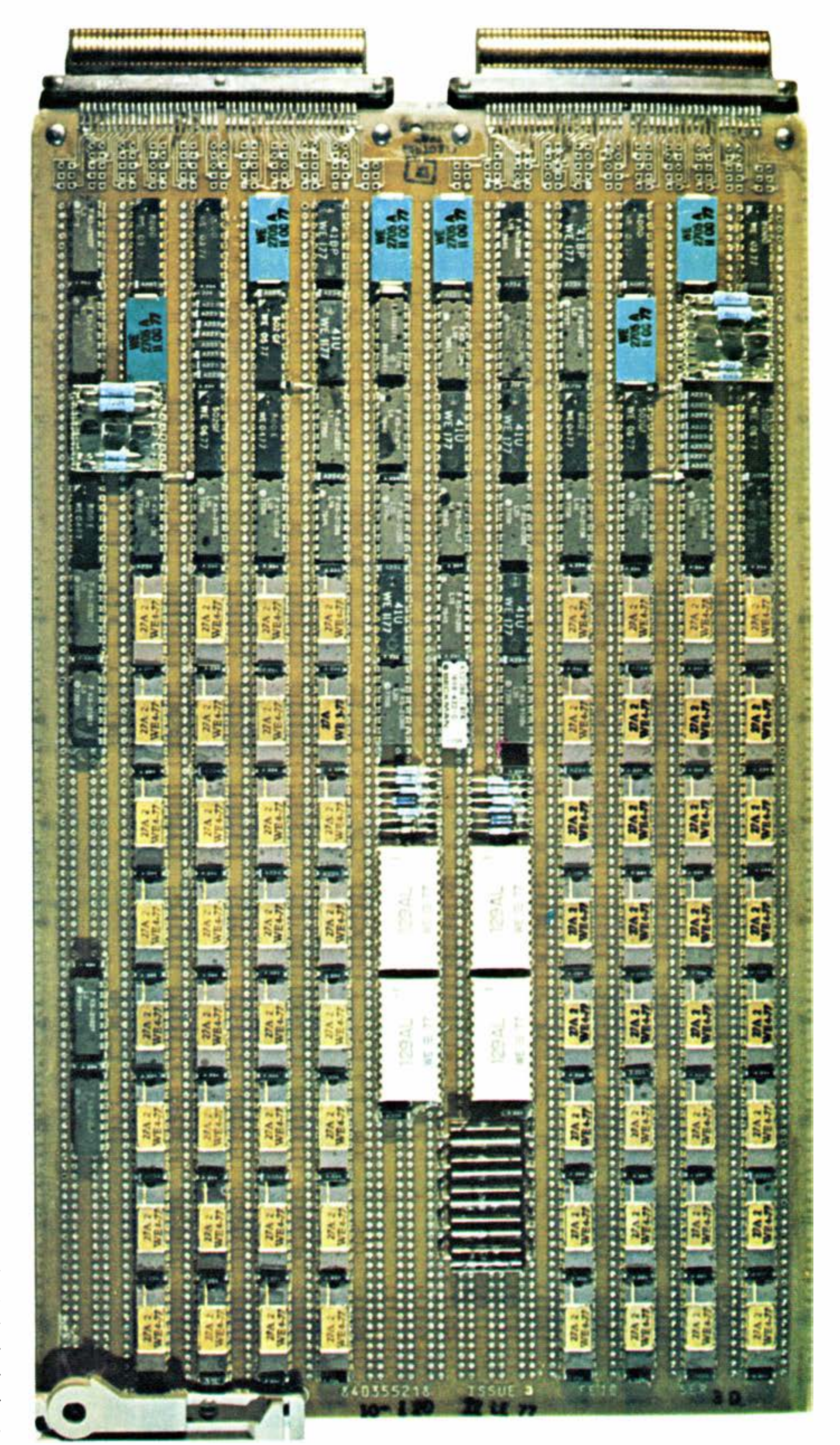
ejemplo. La puerta lógica de tubos de vacío no resultaba particularmente fiable. Una puerta lógica realizada con transistores discretos demostró ser unas 1000 veces más segura que la anterior. Una puerta de un moderno circuito integrado es, por lo menos, 100 veces más fiable que la de transistores discretos. El paso de los tubos de vacío a los circuitos integrados supone, por tanto, un aumento de unas 100.000 veces en la fiabilidad de una puerta lógica.

Las centrales que hacen posible el encaminamiento de las llamadas telefónicas son un ejemplo de lo que significa esta mejora de fiabilidad para un sistema de comunicación. Una máquina de este tipo puede recibir simultáneamente hasta 100.000 líneas telefónicas. Observa las líneas y detecta cuándo un abonado requiere servicio (el momento de descolgar su teléfono). Entonces recoge la información que marca, le conecta con el abonado llamado, registra la información necesaria para el cobro del servicio y desconecta cuando el abonado llamante cuelga.

El "cerebro" de la máquina es un procesador electrónico, que consta fundamentalmente de un control central y de una serie de componentes que almacenan datos. El procesador también realiza otras tareas, incluyéndose en ellas el diagnóstico de fallos y de circuitos estropeados. Puesto que el procesador es crucial para el funcionamiento de los 100.000 teléfonos, debe tener un alto grado de fiabilidad. El objetivo está establecido en que no tenga más de dos horas de avería total en 40 años (la mayor parte de este tiempo se espera que se deba a fallos humanos en la operación y mantenimiento, más que a fallos de los propios componentes electrónicos).

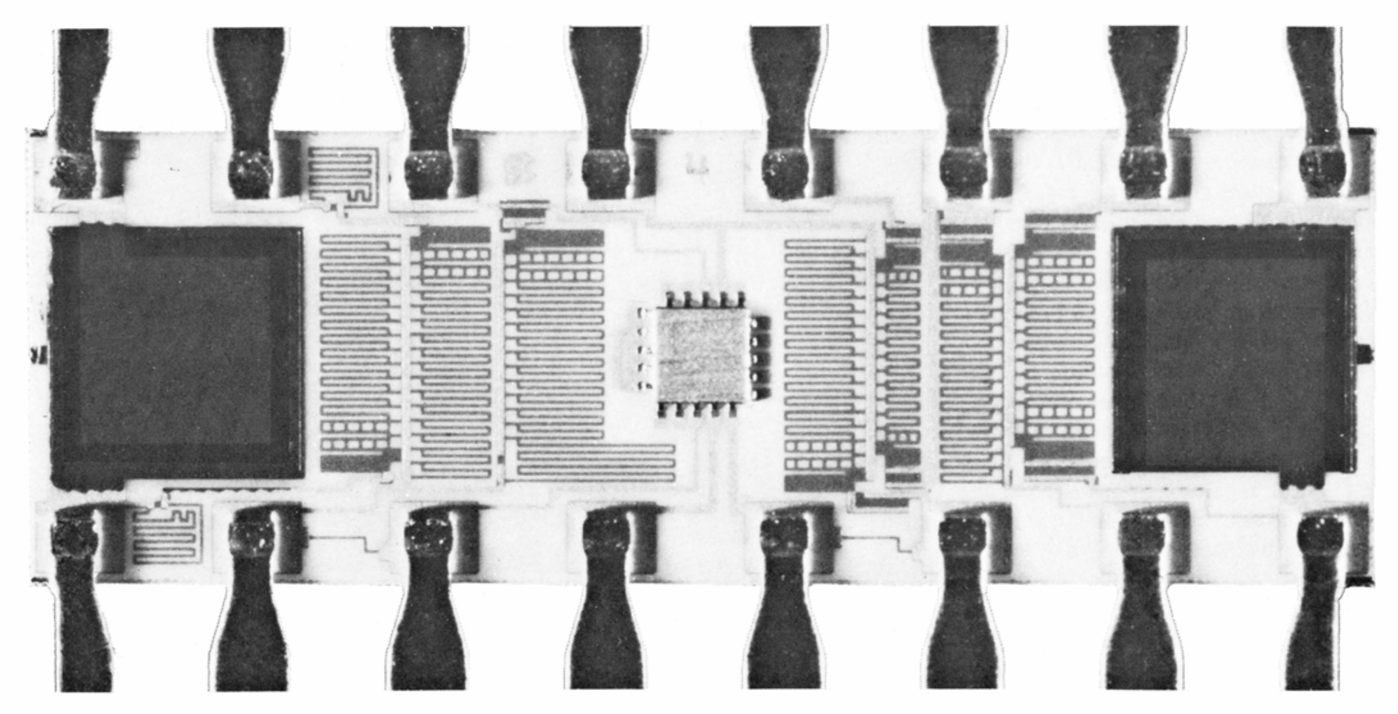
La unidad de medida de la velocidad de fallo de un componente electrónico es el FIT, que equivale a un fallo por cada 10^9 horas de funcionamiento. Considérese el procesador de una gran central. Su control central puede tener 40.000 puertas. Si puede repararse en media hora, podrá tener un fallo cada 10 años, cumpliendo todavía su objetivo. Los fallos de los circuitos son acontecimientos aleatorios, no predecibles fácilmente, existiendo otras causas de fallo en cualquier gran sistema. Tomando un margen razonable para estos últimos factores, el tiempo medio aceptable para el fallo de una puerta debe oscilar en torno a las 10^{11} horas: una velocidad de fallo de aproximadamente 0,01 FIT por puerta.

Esta tasa de fallos no puede alcanzarse fácilmente. Los circuitos integrados a



CONJUNTO DE CIRCUITOS para la unidad de memoria de una central telefónica que ilustra cómo la microelectrónica ha hecho posible la gran disminución de tamaño de los equipos de comunicaciones. El conjunto tiene aproximadamente 33,2 centímetros de largo y 20 de ancho; 161 de estos conjuntos colocados en un bastidor de 0,67 metros de ancho hacen posible la conmutación de hasta 100.000 teléfonos. Apenas hace 15 años, una unidad de memoria similar para una central telefónica electrónica habría necesitado una fila de bastidores de 32 metros de ancho. Aquí los rectángulos dorados son dispositivos de memoria de óxido metálico semiconductor de canal *n* (MOS *n*). La mayoría de los rectángulos de la parte superior del conjunto son decodificadores. Los cuatro rectángulos blancos de la parte central inferior son circuitos integrados híbridos que contienen el programa de la unidad de memoria. La memoria de una central telefónica local almacena no solamente la información permanente necesaria para el funcionamiento controlado por ordenador, sino también información transitoria, como número marcado, duración de llamada y otros.

EL TELEFONO CON MEMORIA, actualmente en servicio en algunas zonas de Estados Unidos, es un resultado de los avances de la microelectrónica. Con los tres botones de la parte superior se puede grabar un número que se solicite frecuentemente. Cada uno de esos números habituales puede marcarse apretando uno de los botones que están debajo de los tres anteriormente citados. El botón de la parte inferior derecha, al lado de la inscripción "Last number dialed" ("último número marcado"), se utiliza cuando se ha escuchado la señal de ocupado. El teléfono "recuerda" el número, por lo que puede llamarse de nuevo (sin necesidad de volverlo a marcar), pulsando solamente el botón arriba mencionado.



FILTRO ELECTRICO DE FRECUENCIA VOCAL. Se trata de uno de los circuitos fundamentales en telecomunicación. Demuestra el efecto de la microelectrónica en hacer más pequeños y más fiables los componentes de telecomunicación. El filtro tiene por misión seleccionar una banda de comunicación vocal y excluir las interferencias de ella, lo cual resulta particularmente importante cuando varias señales vocales están siendo multiplexadas sobre un único par de hilos conductores. El filtro consiste en un amplificador en forma de circuito integrado de silicio *(cuadrado del centro)*, una red de resistencias de precisión realizadas con finas capas metálicas *(líneas sinuosas)* y condensadores de precisión también de capas *(grandes cuadrados a ambos extremos)*. El filtro está fabricado con técnicas de capa fina sobre un substrato cerámico de 1,93 cm de largo y 0,66 cm de ancho. El filtro en forma de circuito integrado sustituye a los caros y voluminosos dispositivos compuestos por grandes bobinas y condensadores que era necesario utilizar para esta función, antes de que hiciera su aparición la técnica microelectrónica.

gran escala se sitúan en su antesala, y tienen además la ventaja del bajo coste; éste permite el uso de circuitos redundantes, que reducen todavía más la probabilidad de avería y, además, protegen de los fallos en otras partes, verbigracia, en el cableado o en las fuentes de alimentación. La redundancia mantiene funcionando al sistema; además, la baja tasa de fallos de las puertas permite que trabajen también las partes redundantes, haciendo posible que el coste de mantenimiento se sitúe a un nivel razonable.

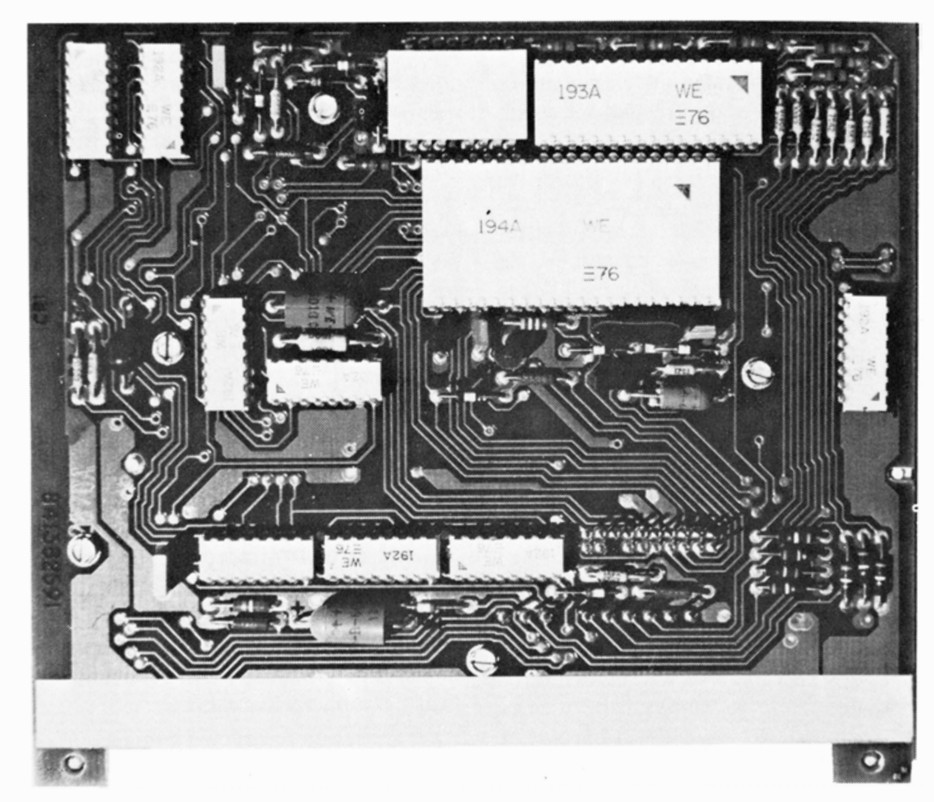
El tamaño del equipo es casi siempre importante en telecomunicación, particularmente en los satélites y naves espaciales. El propio equipo electrónico de tierra cuando es de pequeño tamaño se hace de más fácil manejo y traslado y ocupa menos espacio en los edificios.

Todas estas ventajas conducen a costes menores. La memoria de un procesador de una gran central electrónica puede servir de ejemplo de cómo los avances en microelectrónica han influido en el tamaño.

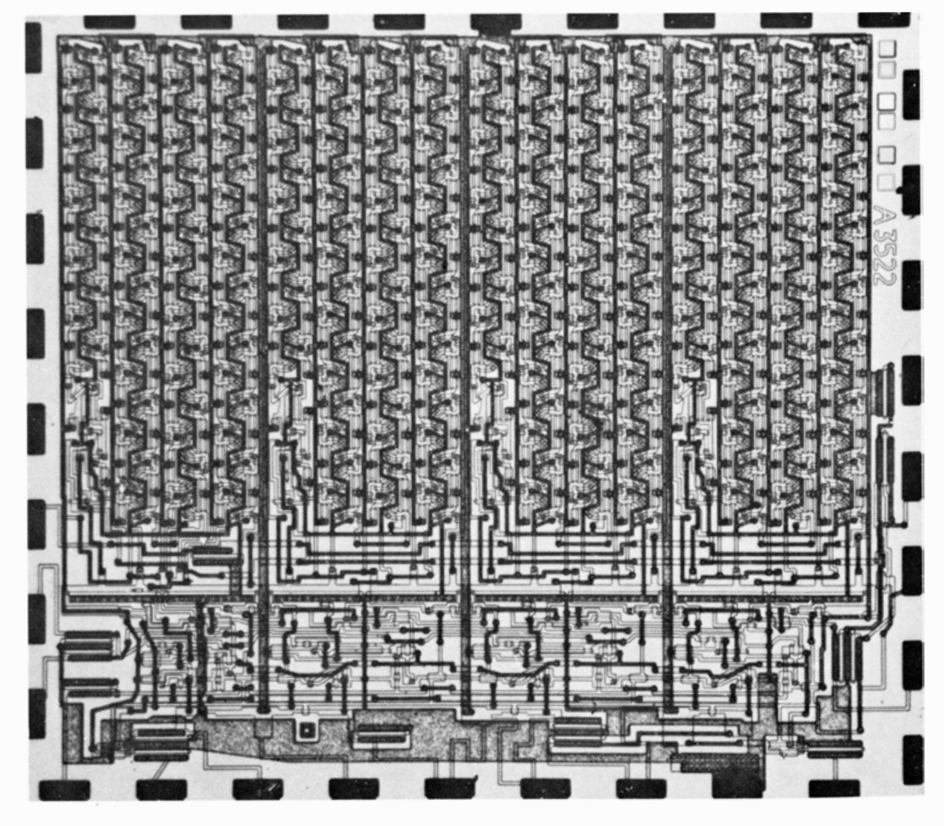
Con la tecnología de los primeros años de la década de los sesenta, la memoria de una central electrónica local necesitaba una línea de bastidores de 31,70 metros de anchura. El almacenamiento de datos se hacía sobre hojas de material magnético. En los años setenta, la misma cantidad de memoria podía ser almacenada en pequeños núcleos toroidales de ferrita, ordenados de forma compacta, que ocupaban una línea de bastidores de unos 2,50 metros. Después aparecieron los circuitos integrados y las memorias de semiconductor, que eran capaces de almacenar primero 1000 bits y después 4000 en una única pastilla de silicio. Utilizando pastillas de memoria de 4000 bits, el circuito de memoria de la misma central podía ser colocado en un único bastidor de 0,67 metros. Ahora que se dispone de circuitos integrados con capacidad para 16.000 bits por pastilla, la misma memoria ocupa solamente un cuarto de bastidor.

Los sistemas de comunicación del futuro usarán, con toda certeza, más circuitos de memoria y de lógica. Su viabilidad sería nula si su memoria necesitase 31,70 metros de bastidores. Resulta ya fácil concebir sistemas que empleen diez veces más memoria que los actuales. La memoria no es más que un botón de muestra de cómo la microelectrónica está posibilitando nuevas ideas en los sistemas de comunicación.

La microelectrónica ha ejercido también un poderoso influjo en los circuitos analógicos, aunque el resultado no ha sido tan espectacular como en el caso de



LA TARJETA DE CIRCUITO IMPRESO del teléfono con memoria lleva a cabo las funciones de registro y "memorización". La tarjeta está montada en la cara posterior de la sección de memoria y está colocada aquí como se vería girando la cara blanca del teléfono, de manera que los tres botones quedaran a la izquierda. Los dos rectángulos blancos de la parte superior izquierda son unidades de memoria, al igual que los otros seis rectángulos semejantes que existen en la placa. Cada unidad es un circuito integrado que puede almacenar cuatro números telefónicos. Los rectángulos blancos mayores, en la parte central derecha, son también circuitos integrados; éstos contienen el programa que ha de desarrollar el sistema y efectúan asimismo las operaciones de marcar.



PASTILLA DE MEMORIA del teléfono "Touchamatic". Aparece tal como se vería abriendo una de las unidades de memoria de la fotografía de la parte superior de esta página. La pastilla está montada normalmente en el centro de un substrato cerámico blanco y contiene los circuitos para almacenar y recuperar los cuatros números telefónicos. Está unida a cuatro de los botones.

las memorias y los circuitos lógicos. Considérese el circuito que quizá sea el más básico en telecomunicación, a saber: el filtro eléctrico de frecuencia vocal, que puede seleccionar una banda de audio y rechazar todas las interferencias de fuera de ella. Estos filtros son particularmente importantes para los sistemas que multiplexan cierto número de canales vocales sobre un mismo hilo conductor; el filtro confina cada canal vocal a cierta banda de frecuencias; en consecuencia, protege a los canales de las interferencias producidas por los otros.

Durante casi 50 años, los filtros se hacían mediante grandes bobinas y condensadores. Entre los años 1920 y 1970 los diseñadores redujeron progresivamente el tamaño de bobinas y condensadores, pero los filtros continuaban siendo voluminosos y caros. En 1969, el progreso registrado en la estabilización de resistencias y condensadores, contra los efectos del tiempo y de la temperatura, y en la fabricación de amplificadores baratos, a base de circuitos integrados, llevó a un filtro equivalente que era algo menor y más barato. El uso de amplificadores integrados (realizados sobre una única pastilla de silicio) permitía al diseñador obtener comportamientos equivalentes a los grandes condensadores y bobinas utilizando condensadores y resistencias de pequeño tamaño. Para el éxito de este propósito fue necesario que los condensadores y las resistencias se mantuvieran altamente estables contra los efectos del tiempo y la temperatura.

Estos requisitos engranaban perfectamente con el desarrollo de microcircuitos en un frente diferente, que implicaban el manejo de películas aislantes y conductoras colocadas sobre substratos cerámicos. A partir de estos trabajos nacieron los circuitos integrados de capa fina, de los cuales ya he hablado. Los circuitos de película realizados con tántalo son, a la vez, baratos y altamente estables. Por cuya razón se emplearon durante algún tiempo como filtros de audio en los circuitos generadores de tono en los teléfonos de teclado, por mencionar solamente una aplicación. Los tonos generados por el circuito deben ser extremadamente estables durante largos períodos de tiempo, puesto que las frecuencias de los tonos distinguen a los dígitos del número marcado.

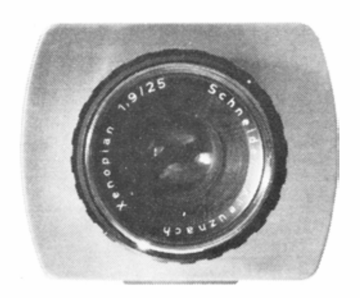
En 1973, esta tecnología había avanzado hasta el punto de que el equivalente al viejo filtro de banda vocal construido con grandes condensadores y bobinas podía fabricarse sobre un pequeño substrato de cerámica. En 1975, el tamaño del substrato se redujo lo suficiente como para hacer que el filtro tenga la apariencia externa de un pequeño componente, muy parecido a una cápsula típica de un circuito integrado. Es de esperar que, muy pronto, la película de tántalo pueda colocarse directamente sobre la superficie de la pastilla de silicio del amplificador. El filtro entero será microscópico. Sin embargo, no concluyen aquí las posibilidades de los microcircuitos en el filtrado de señales analógicas. Los circuitos integrados digitales que trabajan con señales muestreadas y los dispositivos de acoplamiento de carga (CCD, del inglés "charge-coupled devices") presentan grandes posibilidades para continuar el avance de los filtros electrónicos; llegarán a ser incluso más pequeños que los filtros de película fina.

El tamaño importa mucho, sobre todo, en los terminales de comunicaciones. Los teléfonos de todo tipo, las centralitas automáticas (centrales interiores que muchas organizaciones usan para tratar su tráfico telefónico interior o exterior), los terminales que acoplan ordenadores y sensores a los canales de comunicación y las radios de toda clase están a disposición de todo el mundo y ocupan espacio en oficinas, hogares, automóviles y en muchos otros lugares. El pequeño tamaño y otras características interesantes abren nuevas perspectivas para los equipos terminales. Un terminal puede mostrarse asombrosamente potente, siendo pequeño y compacto. No es raro que un terminal de este tipo contenga un microprocesador que, como se ha indicado en este número monográfico, es un pequeño ordenador realizado sobre una pastilla de silicio. El microprocesador de un terminal de comunicaciones cuesta varias decenas de dólares y puede realizar operaciones que diez años atrás hubieran necesitado un gran ordenador, cuyo precio ascendía a cientos de miles de dólares.

Tamaño y peso están estrechamente relacionados. Son ambos de extrema importancia en las comunicaciones espaciales. Es difícil imaginar cómo hubiera podido llevarse a cabo el programa espacial sin el reducido tamaño, el poco peso y la fiabilidad de los circuitos de estado sólido.

Al programa espacial se debe la posibilidad de lanzar satélites de comunicaciones. Cuando un satélite opera con tráfico comercial debe ser competitivo con los sistemas terrestres. El coste es importante y está fuertemente condicionado por el tamaño y el peso de los aparatos electrónicos del satélite. Las exigencias de frecuencia y potencia no permiten que estos aparatos sean realizados enteramente con circuitos integrados, pero los circuitos de este tipo desempeñan un papel notable en la economía



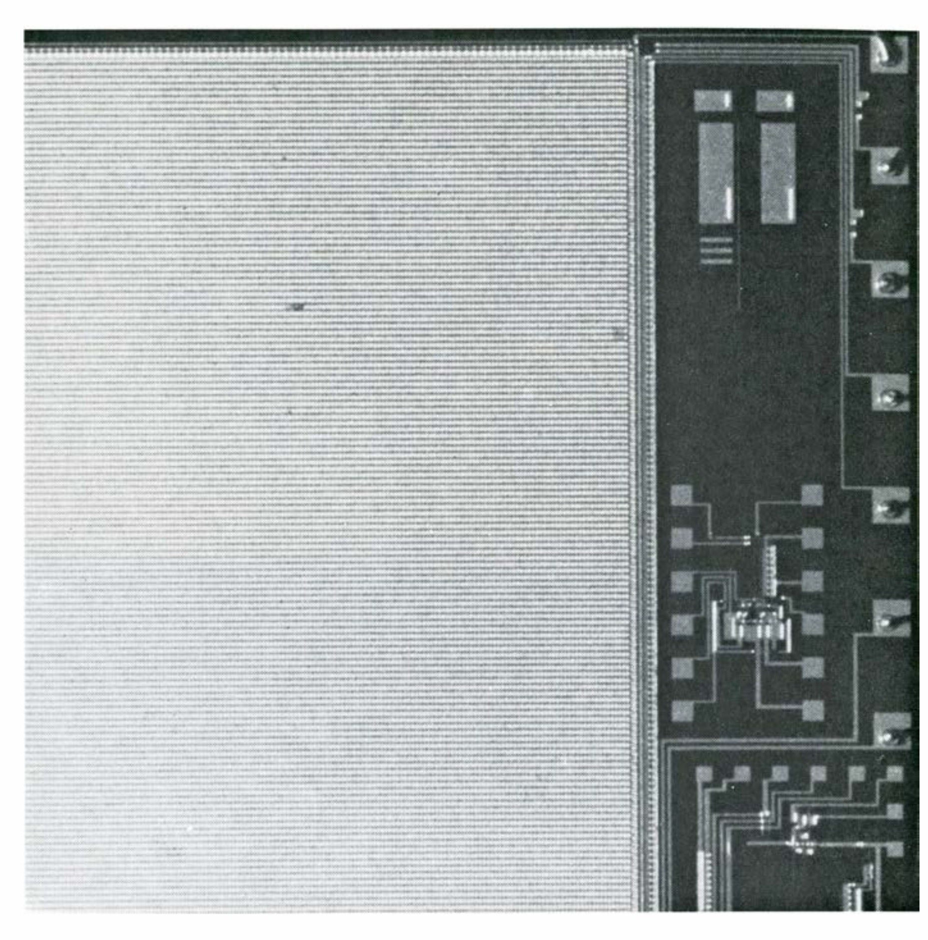


CAMARA DE TELEVISION DE ESTADO SOLIDO, fabricada por Fairchild Camera and Instrument Corporation. Se trata de una muestra representativa del efecto ejercido por la microelectrónica sobre las cámaras de televisión. La unidad clave de esta cámara es un dispositivo de imagen del tipo mostrado en la fotografía de la página siguiente. Gracias a sus componentes de estado sólido, esta cámara es mucho más pequeña que las usuales que emplean el tubo vidicón.

de los sistemas de comunicaciones por satélite. Es más, sin la fiabilidad que ofrecen los componentes de estado sólido habría sido muy difícil llegar a los sistemas de comunicaciones por satélites.

De lo anterior se deduce que el equipo de comunicaciones necesita una buena cantidad de inteligencia propia. Pueden citarse muchos ejemplos, entre ellos los teléfonos que manejan transacciones de caja, terminales de datos que envían información a velocidades prodigiosas, centrales telefónicas con capacidad propia de diagnóstico, centralitas con capacidad ajustada a la necesidad, radios con controles de frecuencia inteligentes, receptores de televisión que se sintonizan solos, equipos de audio de alta fidelidad con especificaciones muy exigentes, etcétera.

Una característica de los circuitos integrados que presenta simultáneamente ventajas e inconvenientes en estas aplicaciones es su limitada capacidad de manejar potencia. Entre las ventajas: el nivel de potencia bastante bajo que necesitan los circuitos de acondicionamiento de señal. Entre los inconvenientes: la mala capacidad de interacción con el ambiente hostil del mundo real. Los circuitos conectados a cables y antenas no son fácilmente protegibles de los rayos, ni los circuitos integrados generan las corrientes de llamada de los teléfonos, ni las fuertes señales de radio necesarias para las comunicaciones a través de la atmósfera.



PORCION DE LA PASTILLA DE IMAGEN para la cámara de televisión de estado sólido; se caracteriza por una fina estructura de sensores formados por dispositivos de carga acoplada, los cuales liberan portadores de carga proporcionalmente a la energía luminosa que les alcanza. La pastilla contiene 185.440 sensores de éstos, de los cuales aproximadamente una cuarta parte resultan visibles en la zona más iluminada de la fotografía, la cual contiene otra circuitería.

La inteligencia no necesita un nivel específico de potencia. Un circuito que cumpla una función de inteligencia puede a menudo trabajar a un nivel de potencia sólo suficiente para llevar los nudos del circuito a tensiones unas pocas veces mayores que las tensiones de cualquier ruido indeseado por el circuito. En los circuitos integrados, la mayor parte del cableado está realizada sobre la pastilla de silicio, y los nudos del circuito son pequeños, necesitándose poca potencia. A medida que la escala de integración ha aumentado, se ha hecho necesario y posible disminuir los niveles de potencia para que los circuitos complejos realizados sobre una única pastilla puedan trabajar a una temperatura aceptable. Sólo unos pocos terminales de entrada y salida que se conectan a otros componentes deben ser atacados con señales suficientemente potentes para entregar señales fuertes.

A menudo, se pueden reducir más los niveles de potencia eligiendo una tecnología que no necesite mucha energía en sus elementos de circuito activos. Una de estas tecnologías es la de los circuitos de estructuras complementarias de óxido metálico semiconductor (CMOS). Un microprocesador que utiliza alrededor de 8000 transistores ha sido diseñado utilizando esta tecnología; ocupa un dado que no llega a 6,50 milímetros de lado. Está específicamente ideado para funcionar como elemento de control de equipos de comunicación. La pastilla puede realizar 434 instrucciones diferentes, trabaja a velocidades de hasta dos megahertzios (dos millones de ciclos por segundo) y consume menos de 0,1 vatios de potencia. Un sistema que ofrece esta cantidad de inteligencia con tan bajo consumo permite que los equipos de comunicaciones funcionen incluso donde no existen líneas de distribución de energía o son poco fiables. Hace posible que los equipos complejos remotos, conectados a largas líneas o cables, se alimenten con la potencia que les llega por los mismos conductores que las señales.

El progreso de la microelectrónica ha sido más lento en las áreas de las comunicaciones que necesitan alta potencia. La capacidad de manejar potencia de los circuitos está limitada por la máxima temperatura a la que la pastilla de semiconductor puede trabajar. Una pastilla que contenga muchos transistores se calentará más que otra que tenga sólo uno.

Los dispositivos de estado sólido pueden emplearse en los circuitos de radiofrecuencia de reducida potencia, pero generalmente deberán ser transistores discretos. Los equipos de radio y de registro y reproducción han reducido su coste y tamaño gracias a los amplificadores integrados de potencia de sus etapas de audio. Los conmutadores de estado sólido de reducida potencia empiezan a sustituir a los conmutadores electromecánicos en los grandes sistemas de comunicación. En los equipos de conversión de energía, los componentes de estado sólido son de creciente importancia en el aumento de eficiencia y en el abaratamiento de los costes.

El apremio por extender la capacidad de la microelectrónica a los circuitos de comunicación de alta potencia es insistente. Se necesitarían circuitos capaces de soportar cientos de voltios para proteger de los rayos a los sistemas conectados a cables, y para transmitir la relativamente moderada potencia capaz de

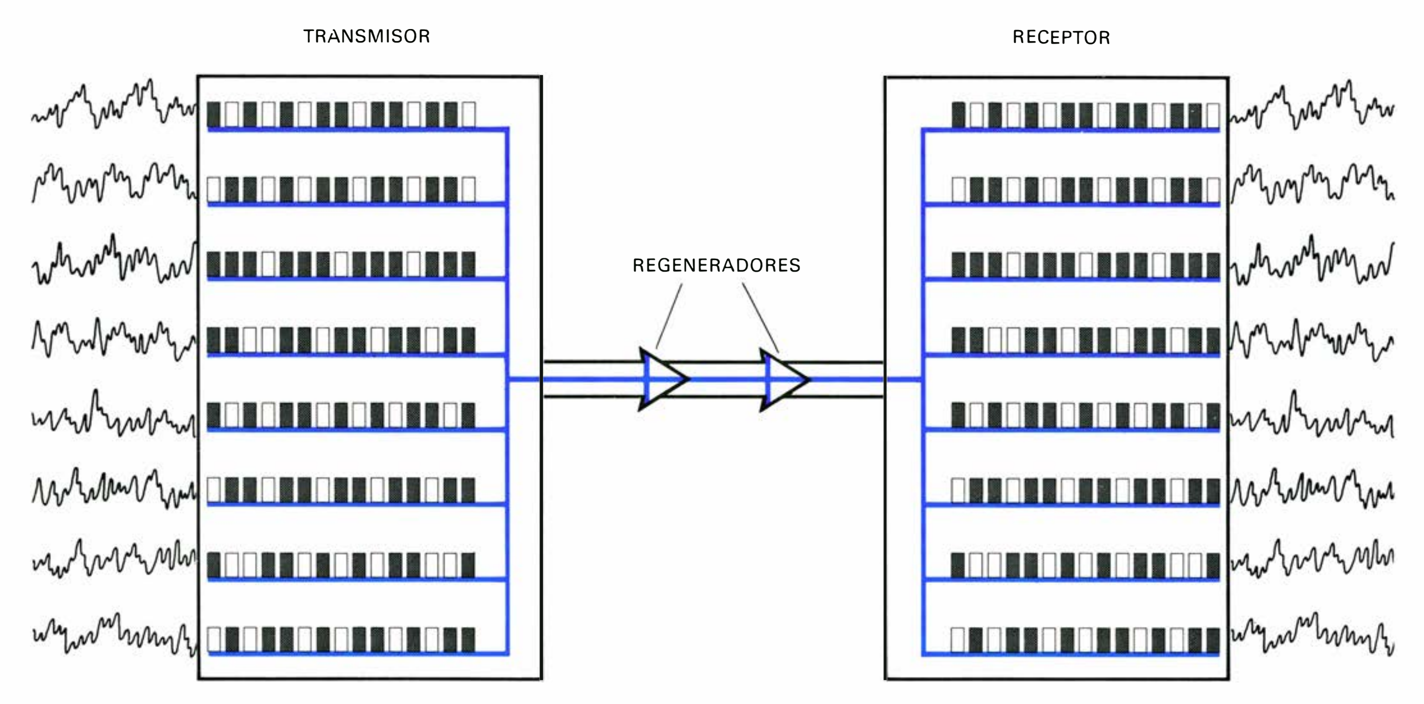
excitar el timbre de un teléfono. No es probable que los circuitos integrados sean alguna vez útiles para estos fines, en las escalas de integración ahora encontradas en los circuitos capaces de manejar inteligencia. Las altas tensiones pueden destrozar cualquier material de dimensiones suficientemente pequeñas y pueden, por otra parte, reducir la fiabilidad del dispositivo. Un circuito de alta tensión debe tener, pues, sus elementos convenientemente separados.

Muchas de las necesidades de las técnicas de comunicación pueden ser mejor satisfechas con las técnicas digitales. Las señales transmitidas en forma de dígitos no se degradan, siempre que los dígitos se reciban correctamente. Una gran cantidad de ruido o diafonía, que sería rechazable para una señal analógica, del tipo de las usadas para transmitir la voz humana, puede no tener ningún efecto degradante sobre las señales digitales. Siempre que el ruido y la diafonía sean algo más débiles que la señal digital, puede detectarse la presencia de cada dígito binario y regenerarse completamente la señal digital. La señal se limpia así de ruido y diafonía, lo cual supone una gran ventaja cuando se transmite la señal a través de un medio ruidoso.

Un sistema de comunicación digital debe trabajar a alta velocidad de impulsos y debe emplear gran número de puertas lógicas. Estos circuitos son pequeños y bien adaptados a las modernas tecnologías de la microelectrónica. Desde 1962, las señales de voz se han transmitido en forma de dígitos en parte de la red telefónica de Estados Unidos. Las transmisiones emplean modulación por impulsos codificados, en la que la voz es muestreada repetidamente a una alta velocidad y las muestras quedan codificadas en forma de trenes de números binarios. Los números binarios se transmiten en forma de señales digitales, que se van regenerando periódicamente a lo largo de



MODULADOR-DEMODULADOR, fabricado por la Bell System; sirve para transmitir datos digitales sobre un circuito analógico del tipo de los que transportan la mayor parte del tráfico vocal en la red telefónica estadounidense. El aparato se denomina modem (contracción de modulador-demodulador) o equipo de datos. En el extremo receptor hay un instrumento similar que convierte las señales de nuevo en digitales.

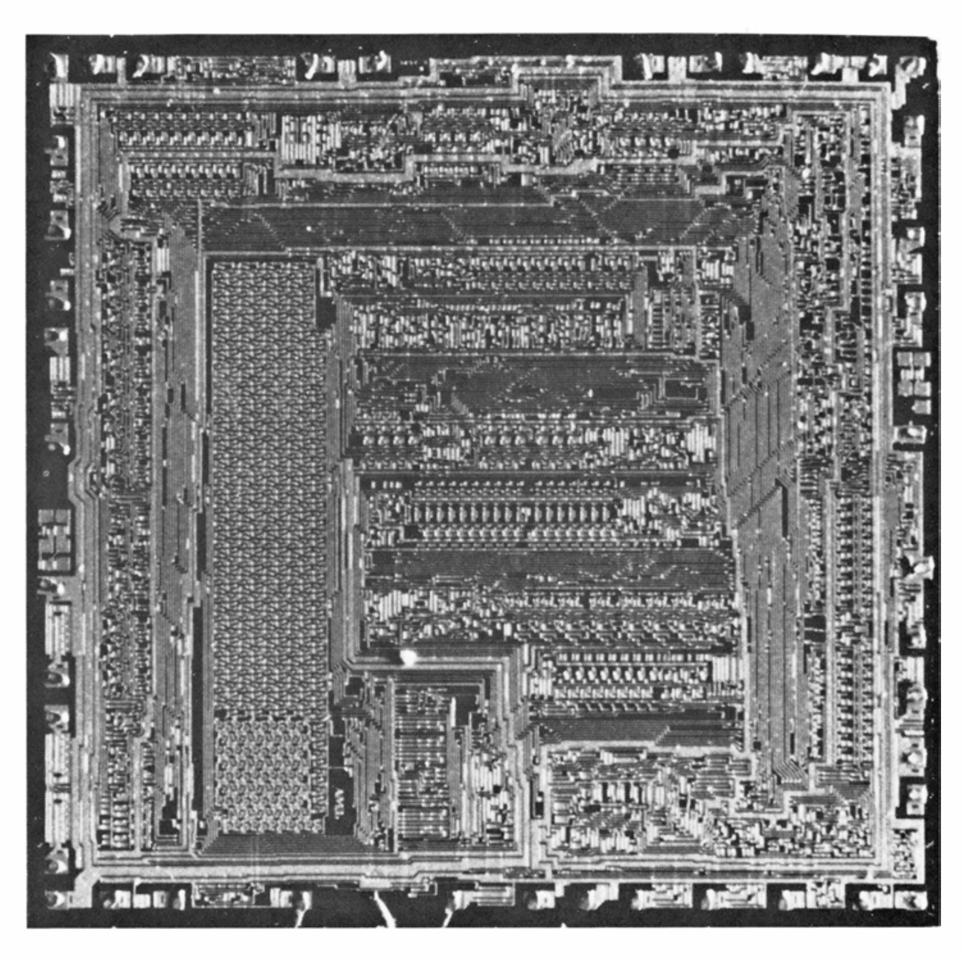


LA TRANSMISION DIGITAL de las señales vocales se hace de forma creciente en los sistemas telefónicos estadounidenses, como resultado del desarrollo de la microelectrónica. Las ondas de sonido *(izquierda)* generadas por la voz son muestreadas a alta velocidad y las muestras convertidas en señales digitales, que se transmiten sobre las líneas telefónicas como pulsos presentes *(blancos)* o pulsos ausentes *(negros)*. Periódicamente, las señales se van regenerando a lo largo de la transmisión. En el extremo receptor, las señales digitales se reconvierten en señales de voz. Los microcircuitos desempeñan un papel importante no sólo en lo referente a la conversión de la palabra en dígitos sino también en la multiplexación de un gran número de conversaciones sobre un cable y en la regeneración de las señales durante la fase de transmisión.

la ruta para extraer de ellas el ruido, la diafonía y la distorsión. En el extremo receptor, los números binarios se convierten de nuevo en muestras de señal, a partir de las cuales puede reconstituirse el sonido de la voz.

Las ventajas de la comunicación digital se conocían desde hace algún tiempo. La extensión de su aplicación ha venido condicionada fundamentalmente por la economía relativa de las circuiterías analógicas y digitales. El desarrollo de los circuitos integrados ha cambiado en gran medida la situación, inclinando la balanza en favor de las técnicas digitales en muchas aplicaciones.

Un inconveniente de las señales digitales es que precisan, para su transmisión, mayores anchos de banda que las señales analógicas. Por tanto, las técnicas analógicas continúan siendo atractivas en medios apantallados de ancho de banda limitado, tales como los cables coaxiales, y en propagación por radio, en donde el ancho de banda está limitado por el espectro disponible. Muchos esfuerzos, entre los que cabe incluir gran variedad de enfoques, se han dedicado a reducir la velocidad de los impulsos necesaria para transmitir la voz digitalmente. El bajo coste de la microelectrónica es esencial en casi todos ellos. La velocidad de impulsos por canal de voz va desde 2400 bits por segundo, en sistemas muy complejos pero de baja calidad, hasta unos 64.000 bits por segundo, en los sistemas de modulación por impulsos codificados de alta calidad. Las tasas más bajas introducen degradaciones de la señal, que pueden o no ser aceptables para una aplicación dada.



LA PASTILLA DEL MODEM es la parte central del funcionamiento del modulador-demodulador. La pastilla es un circuito integrado a gran escala que convierte la señal digital o la onda cuadrada, procedente de un ordenador o de otras máquinas de proceso de datos, en el formato de onda modulada necesario para la mayoría de las líneas telefónicas norteamericanas. Dentro de unos años, las líneas telefónicas se modificarán y podrán transportar las conversaciones en forma de señales digitales.

Existe hoy un enfoque de particular interés, a saber: la interpolación de la conversación por asignación de tiempo (TASI, del inglés "time - assignment - speech - interpolation") idea originalmente desarrollada por los laboratorios Bell para circuitos de cables submarinos que transportan señales vocales por técnicas analógicas. La idea consiste en aprovechar los breves silencios normales de las conversaciones para retirar el canal de comunicación de un abonado momentáneamente callado para asignarlo a otro que está hablando. El malabarismo es tan veloz y preciso que los abonados no se dan cuenta de ello. La idea puede ser ahora materializada usando microcircuitos y técnicas digitales, convirtiendo un gran número de señales vocales en dígitos,de suerte que se consigue reducir bastante la velocidad de bit efectivamente asignada a cada canal de voz.

Otra idea que se ha visto favorecida por la microelectrónica es la de control por programa almacenado. Primero mediante potentes ordenadores asociados a grandes equipos, como en el caso de las centrales, y más recientemente con microprocesadores incluidos en cualquier tipo de aparatos, la microelectrónica ha aportado a las comunicaciones un conjunto rápidamente creciente de dispositivos programados. Las características de servicio y funcionamiento de muchos grandes sistemas pueden hoy alterarse reescribiendo o añadiendo sentencias a los programas almacenados en la memoria.

Hasta hace poco, un sistema de comunicaciones debía ser recableado o sustituido para cambiar el tipo de servicio ofrecido. Mediante control por programa almacenado, el viejo soporte físico puede programarse para realizar nuevas funciones; éstas tienen que ver más con un mejor aprovechamiento del personal que opera y mantiene el sistema que con las características de los servicios que afectan a los abonados. Los sistemas son casi siempre digitales, empleando una gran cantidad de lógica y memoria. La microelectrónica ha extendido ampliamente los horizontes de las técnicas de control por programa almacenado al ofrecer circuitos lógicos y de memoria baratos y altamente fiables.

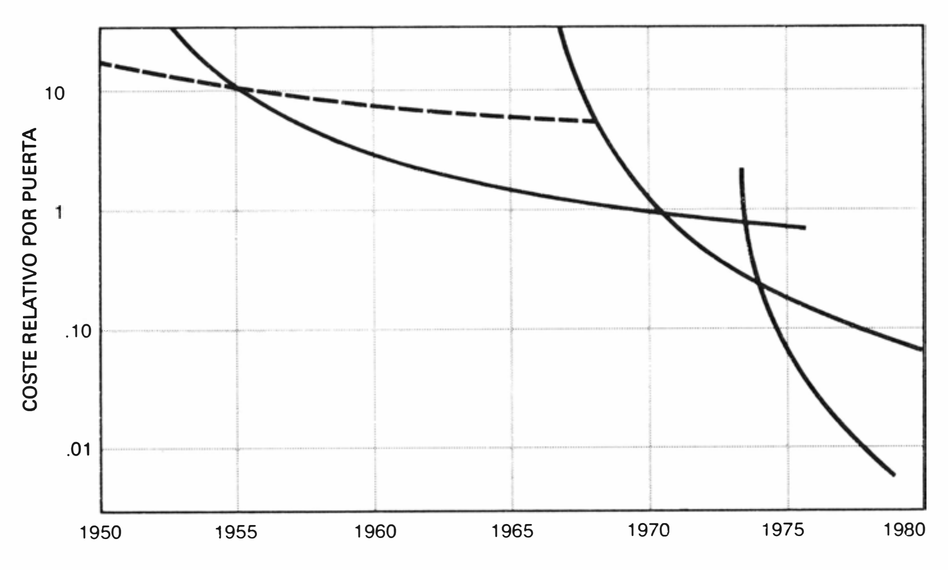
Aunque el control por programa almacenado es una idea simple, el desarrollo de un programa que esté libre de errores es difícil. No es ninguna rareza que el desarrollo de un programa para un pequeño microprocesador del tipo usado en terminales de comunicaciones llegue a costar 100.000 dólares. El coste por el desarrollo de un programa para equipos grandes, como en el caso de una central, puede valer varios millones de dólares.

El silicio ha sido para la microelectrónica lo que el acero para la revolución industrial. Rara vez, sin embargo, un sistema de comunicaciones puede realizarse completamente con circuitos de silicio. Hay otras tecnologías de la microelectrónica que resultan de importancia para las comunicaciones.

Por ejemplo, cierto número de materiales semiconductores (arseniuro de galio y fosfuro de galio entre ellos), emiten luz cuando conducen corriente. Los

diodos emisores de luz se usan ampliamente como indicadores e iluminadores. Presentan también los números y las letras en las lecturas de las señales digitales. Tales dispositivos tienen enorme interés en las comunicaciones porque permiten pasar las señales eléctricas a una información entendible por el cerebro humano, sin necesidad de recurrir al papel escrito.

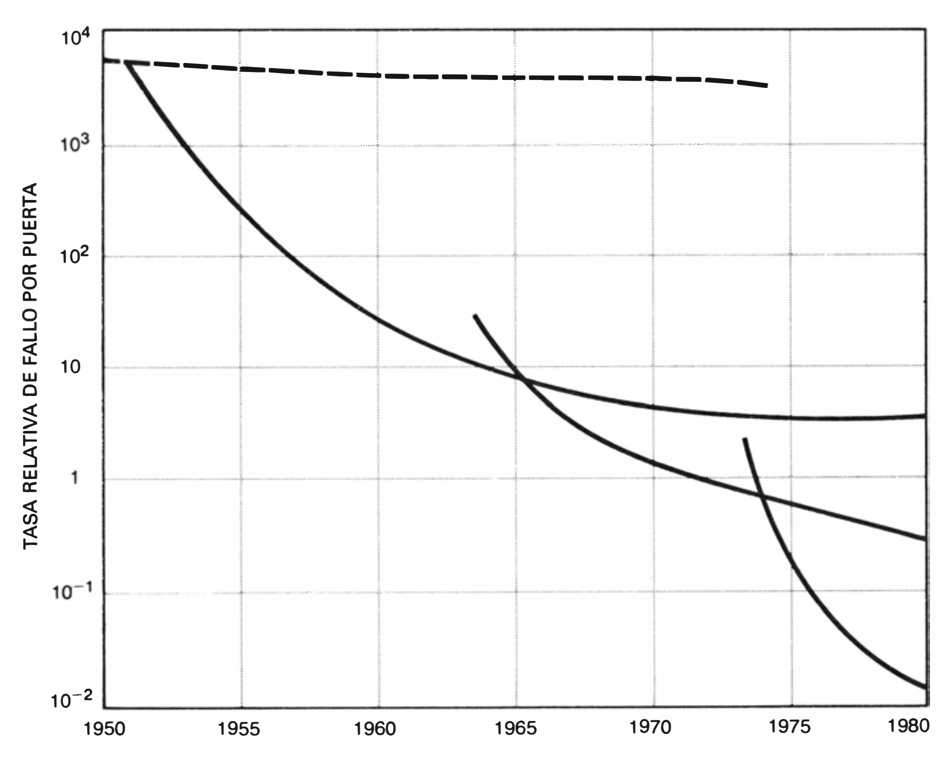
De mayor importancia potencial en el campo de las comunicaciones llega a ser el láser de estado sólido. Se trata de un sofisticado primo de los diodos emisores de luz, al ser esencialmente un diodo de estructura compleja que emite un rayo de luz altamente colimado y monocromático. Su utilidad se apreciará sobre todo en los sistemas de comunicaciones ópticas. Los problemas se han centrado en la fiabilidad de los láseres. Hace sólo cinco años, los mejores láseres de estado sólido duraban sólo unas pocas horas funcionando. Ahora, puede conseguirse que esos láseres funcionen durante un millón de horas, lo cual es suficiente para hacer de este dispositivo un atractivo elemento para las redes de telecomunicación.



LA REDUCCION DEL COSTE de los componentes de los sistemas de comunicaciones es uno de los efectos de la microelectrónica. Se ilustra aquí para las puertas lógicas, que son los elementos clave de los sistemas de comunicaciones y de cálculo. Las curvas representan respectivamente puertas de tubos de vacío, puertas de transistores discretos, puertas de pequeña escala de integración (unas pocas puertas por pastilla) y puertas de gran escala de integración (miles de puertas por pastilla). En cada caso, el dispositivo representado es una puerta lógica digital de la mejor calidad.

Los circuitos de burbujas magnéticas constituyen otro producto de gran porvenir en la investigación de las comunicaciones. Las burbujas magnéticas son pequeños dominios magnéticos cilíndricos que pueden existir y moverse en ciertos materiales. Se ha fabricado granate de composición especial que puede contener burbujas magnéticas de pocas micras de diámetro. Realmente, la mayor escala de integración jamás lograda ha sido en las memorias de burbujas magnéticas. Se suelen fabricar pastillas de granate de hasta 64.000 bits de memoria por pastilla y de forma experimental se han construido algunas de 256.000 bits. Dispositivos de burbujas magnéticas con cuatro pastillas de 64.000 bits por cápsula están actualmente prestando servicio en la red telefónica, formando parte de una máquina, que no tiene partes móviles y que suministra avisos previamente grabados. Las burbujas magnéticas tienen la importante característica de no desaparecer cuando se interrumpe la alimentación del circuito.

Una estructura funcionalmente similar realizada en semiconductores es el dispositivo de carga acoplada. Pueden transferir una cadena de cargas eléctricas a través de un semiconductor de una forma similar al movimiento de un tren de burbujas magnéticas en un granate. Desgraciadamente, la carga en uno de estos dispositivos desaparece cuando se interrumpe la alimentación. Por el contrario, tiene la ventaja de realizarse con las técnicas fotolitográficas y los procesos que se siguen en la fabricación de circuitos integrados. Se trata de un circuito integrado de configuración especial. Ofrece la ventaja adicional de la alta movilidad de los portadores de cargas, dando lugar a dispositivos que son muchas veces más rápidos que los de burbuja magnética. La tecnología de los dispositivos de carga acoplada ha sido aplicada a la cámara de televisión de estado sólido y en dispositivos que necesitan memoria en serie. Su empleo se muestra asimismo prometedor en los filtros electrónicos de los sistemas de comunicaciones.



LA GANANCIA EN FIABILIDAD de los componentes electrónicos es otro de los efectos más importantes de la microelectrónica en las comunicaciones. Las dos mejoras juntas (reducción de costes y aumento de la fiabilidad) han dejado sentir su peso en las comunicaciones. En este diagrama, las curvas representan puertas lógicas de alta calidad de tubos de vacío, de transistores discretos, de pequeña escala de integración y de gran escala de integración, respectivamente.

Microelectrónica e informática

La integración a gran escala permite que los elementos lógicos sean más rápidos y más baratos, aunque el movimiento de datos permanece lento y caro. Se requieren nuevos diseños y teorías, basados en el proceso en paralelo y en la regularidad geométrica

Ivan E. Sutherland y Carver A. Mead

La informática se ha desarrollado en una era tecnológica en la que los cables eran baratos y los elementos de conmutación caros. La tecnología de los circuitos integrados está invirtiendo esta situación económica, haciendo que los elementos de conmutación sean prácticamente gratuitos y que los cables sean el único componente caro.

Dentro de un circuito integrado, "los cables", que son los caminos conductivos, resultan caros porque ocupan la mayor parte del espacio y consumen la mayor parte del tiempo. Entre circuitos integrados, "los cables", que ahora suelen ser las pistas conductoras de un circuito impreso, son caros por su tamaño y por el efecto de frenado de las señales. La informática empieza ya a tener en cuenta esta inversión de costes. El diseño de los ordenadores todavía no saca pleno partido de la completa gama de posibilidades implícita en la microelectrónica. Cuando aprendamos a comprender la inversión de costes entre lógica y cables y a aprovechar las posibilidades inherentes a la integración a gran escala, será cuando podremos esperar que se produzca una revolución real en la práctica del cálculo, no sólo en la forma de las máquinas, sino también en las teorías en las que se fundamenta su diseño y uso.

¿Por qué deben revisarse las teorías de la computación? Supongamos que nos dediquemos a desarrollar algunas teorías de computación, en la esperanza de que nos sirvan para conseguir dos fines: establecer cotas superiores a lo computable y servir de pauta para el diseño y el empleo de las máquinas de cálculo. Es de presumir que tales teorías harían avanzar el conocimiento de los procesos de computación y quizás vertieran luz sobre la naturaleza del conocimiento y del pensamiento. Las teorías pueden basarse únicamente en el razonamiento matemático o pueden apoyarse también en los principios fundamentales de la física.

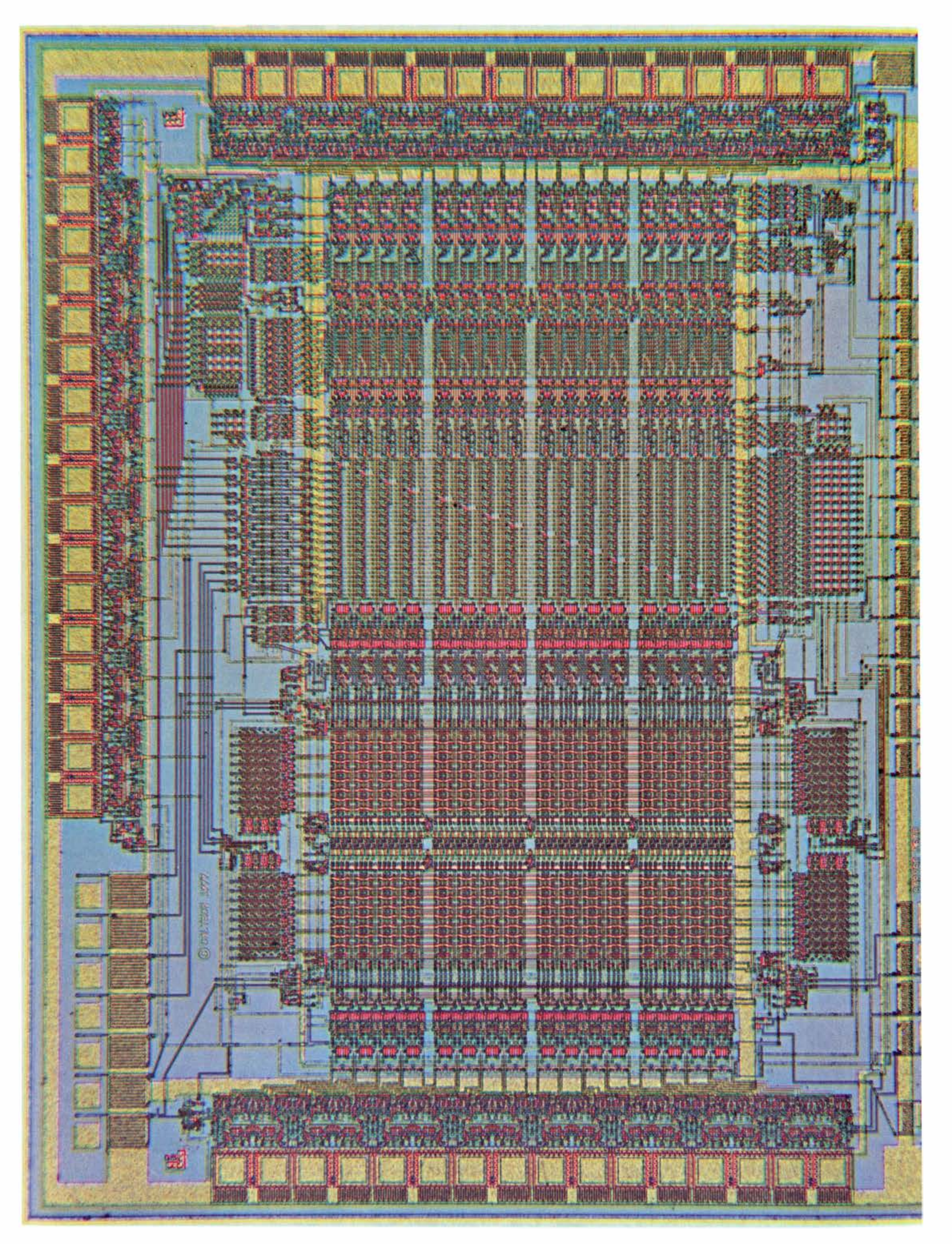
De las que se basan solamente en el razonamiento matemático podemos esperar que nos sirvan para demostrar propiedades de los ordenadores que no tengan relación con los principios físicos. Sólo teniendo en cuenta los principios físicos, podemos, sin embargo, determinar propiedades cuantitativas sobre cuánto tiempo empleará un ordenador de determinado tamaño en realizar un determinado proceso; esto se debe a que la información sólo puede viajar dentro del ordenador a velocidad inferior a la de la luz, y por otro lado a que se necesita cierta cantidad de materia, energía y espacio para representar un bit, es decir, un dígito binario de información, con determinada fiabilidad.

La informática, tal como se la entiende hoy día, se basa casi exclusivamente sobre razonamientos matemáticos. Atiende a las operaciones lógicas que tienen lugar dentro de los dispositivos de cálculo. Sólo de forma superficial considera la necesidad de distribuir los dispositivos lógicos en un determinado espacio, y la obligación subsidiaria de establecer enlaces de conexión entre ellos. La informática, hoy día, tiene poco que decir acerca de cómo las restricciones físicas de tales conexiones limitan la complejidad de las tareas de cálculo realizables "de facto" en un ordenador determinado.

Es así, en parte, porque quien concibe un ordenador como una máquina lógica que hace operaciones lógicas, numéricas y algebraicas sobre datos, tiene cierta inclinación a pensar en la máquina en términos de notaciones matemáticas relativas a estos campos. En tales notaciones, un símbolo x escrito en un lugar de una página significa exactamente lo mismo que el símbolo x escrito en otro lugar de la página. La idea de que hace falta una conexión espacial, para conseguir que ambos valores representen lo mismo en un dispositivo de memoria de calculador, no aparece en la notación. La notación, en sí misma, dirige la atención hacia las operaciones lógicas, reflejando el hecho de que los seres humanos sólo piensan, de modo efectivo, en una sola cosa cada vez. Una prueba matemática es una secuencia de pasos que realizamos a lo largo de determinado período de tiempo; y resulta fácil pensar que los dispositivos de cálculo sólo hacen también una única cosa a la vez. Este enfoque secuencial matemático no es esencial al ordenador, pero el planteo matemático que normalmente damos a los problemas no anima, precisamente, a pensar en enfoques distintos de los secuenciales para la solución de dichos problemas. Casi todos los ordenadores que operan hoy en día hacen pasos individuales de cálculo sobre elementos individuales de datos, uno tras otro, en secuencia temporal.

Era sensato no prestar atención al coste de las conexiones en la época en que los elementos lógicos eran caros y lentos y los cables, por el contrario, rela-

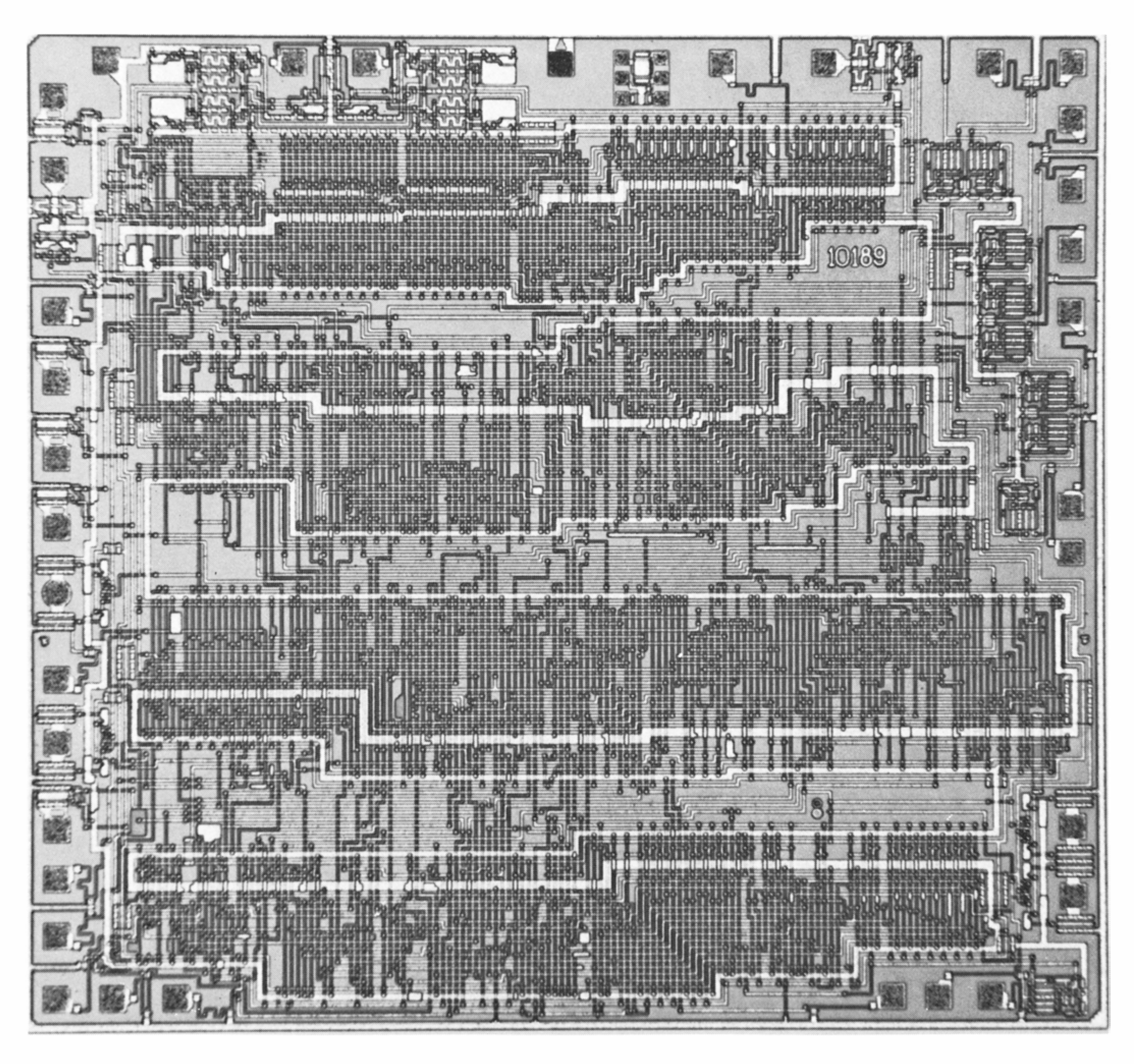
CIRCUITO "OM", microprocesador experimental diseñado por los autores de este artículo en el Instituto de Tecnología de California. Se caracteriza por su alto grado de regularidad, que hace posible más funciones lógicas y de memoria en una sola pastilla. El núcleo de la pastilla (prescindiendo de las uniones de comunicaciones de la parte superior e inferior) está formado por dieciséis columnas prácticamente idénticas, dispuestas en cuatro grupos de cuatro; cada columna representa un bit en un calculador de palabra de 16 bits. Cerca del 40 % de la pastilla *(parte inferior)* **es la memoria; el 20 % central, la sección de decalaje y el 20 % superior, la sección aritmética.**



tivamente baratos y rápidos. Las máquinas secuenciales se adaptaban bien a este tipo de tecnología, porque se podían construir con un número mínimo de elementos de conmutación. Se ha llegado –sea por natural inclinación, sea por influencia de las notaciones matemáticas y tecnológicas– al desarrollo de un tipo de máquinas de cálculo y un cuerpo de teorías de computación, que tanto unas como otras, han quedado pasadas de moda al aparecer la tecnología de los circuitos integrados. Hemos sido capaces de no prestar atención a las limitaciones derivadas de los principios físicos de las conexiones internas de los ordenadores, porque dichas comunicaciones no influían notablemente en los tiempos de cálculo, y eran sólo una pequeña parte de los costes de las máquinas que construíamos. Pero ahora que los elementos lógicos son prácticamente gratuitos y ha quedado el coste de las conexiones como factor dominante, la tecnología de los circuitos integrados nos obliga a una remodelación, no sólo en los tipos de máquinas que se construyan, sino también en sus fundamentos teóricos.

No será fácil desarrollar una nueva teoría para la informática. De hecho, la tarea ha sido ya en parte negligida porque es muy difícil combinar nociones de lógica con nociones de topología, de tiempo, de espacio y de distancia, tal como la nueva teoría requiere. En este artículo vamos a exponer algunos de los elementos que tal teoría debe incluir. Primero, examinaremos los defectos de una teoría sencilla aplicable a redes lógicas pequeñas. Más tarde, veremos cómo los cambios de los costes relativos de lógica y cableado han de influir en la estructura de los ordenadores que se construyan en el futuro. Finalmente, confiamos en describir algunos de los elementos que creemos que deben figurar en una teoría de computación adecuada a las nuevas estructuras.

Una teoría de este tipo será distinta de la que actualmente sirve de base a la informática, y nos parece justificado



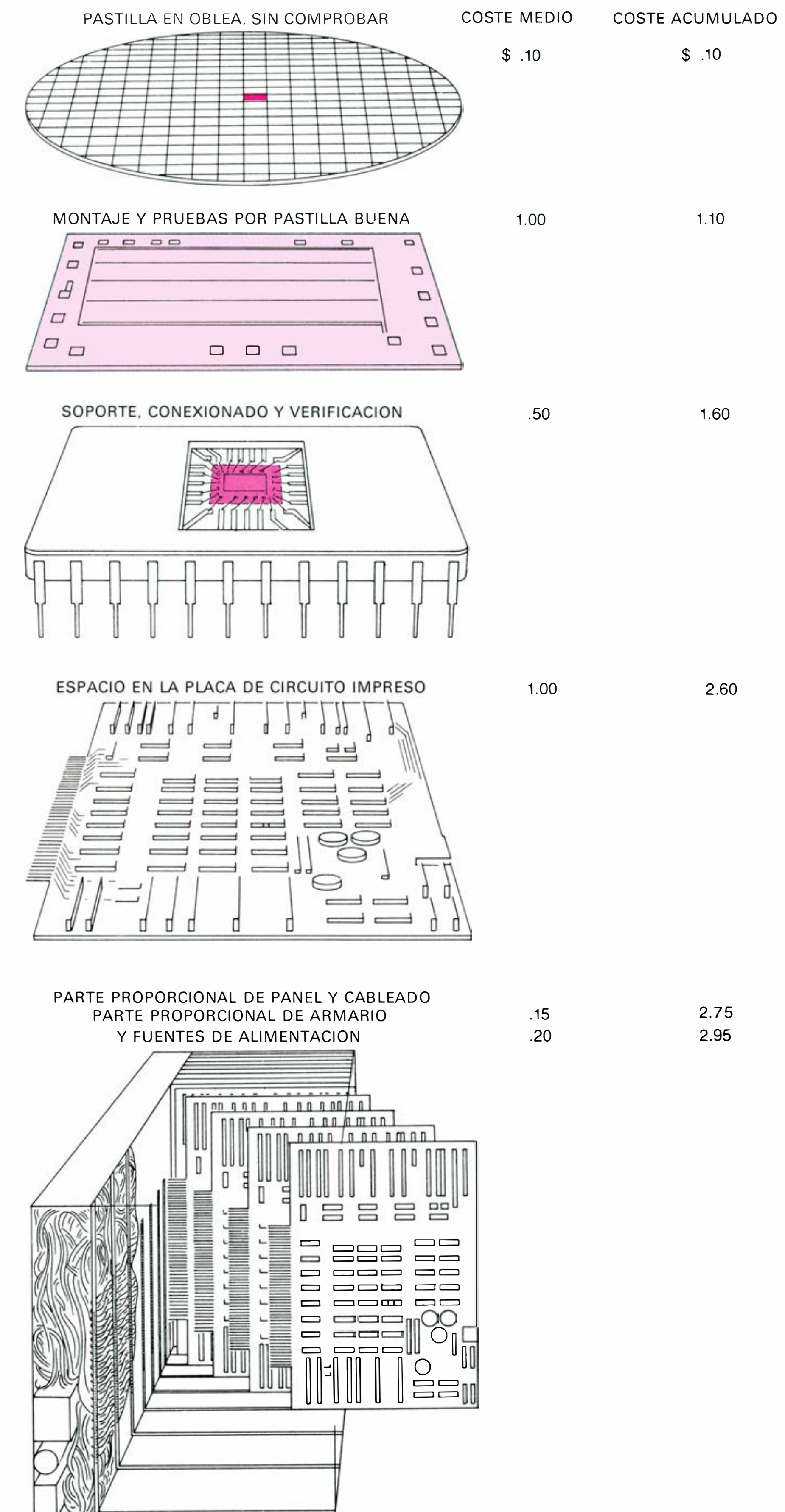
LAS INTERCONEXIONES entre elementos lógicos de un circuito integrado resultan más caras que los mismos elementos; es más, se hacen más largas y más numerosas a medida que incrementa la complejidad de las matrices de elementos cableados aleatoriamente. Aun en niveles de complejidad modesta, el cableado ocupa la mayoría del espacio disponible en una pastilla de circuito integrado. Este circuito integrado, relativamente simple, data de 1971. Nótese que los conectores lineales situados entre los elementos activos, ocupan la mayor parte del espacio.

que describamos como revolucionario el efecto de la tecnología de circuitos integrados, no sólo en el diseño de las máquinas de cálculo, sino también en los esquemas intelectuales que este tipo de máquinas ha utilizado.

La mayoría de las facultades de informática incluyen en sus programas de estudio cursos de teoría de la conmutación; lo que no tiene demasiado valor para la práctica actual del diseño de ordenadores. La teoría de la conmutación, que se inventó para servir de ayuda al diseño de los circuitos de conmutación de redes de sistemas de teléfono automático, aportaba pautas al diseñador que tenía que crear circuitos con un número mínimo de relés para conseguir determinada operación lógica. Luego se extendió al diseño de redes que incluyesen otro tipo de elementos lógicos, por ejemplo, redes lógicas que tuviesen un número mínimo de puertas lógicas de tipo tradicional.

Sin embargo, no hay ninguna garantía de que una tal red con un número mínimo de elementos lógicos ocupen un espacio mínimo en un circuito integrado actual, o realicen su trabajo en un tiempo también mínimo. Los diseñadores de circuitos integrados se encuentran a menudo con que pueden añadir transistores a un esquema y, a pesar de todo, ganar espacio y tiempo, porque añadiendo elementos al número mínimo pueden simplificar la trama de conexiones que hacen falta en el diseño, lo que en definitiva puede contribuir a acelerar la operación. La teoría de la conmutación, hay que reconocerlo, minimiza el número de elementos de conmutación, pero no toma en cuenta el coste y el retraso de los enlaces. En la tecnología actual, el área de un circuito dedicada a conexiones entre elementos suele ser muy superior al área ocupada por los elementos de conmutación, y los retrasos de comunicación mayores que los provocados por los circuitos lógicos. Por tanto, lo que se necesita es una nueva teoría que minimice el coste de las tareas de computación, considerando, no sólo el coste en superficie y tiempo de los elementos de conmutación, sino también el tiempo y área gastados en el proceso de trasladar los datos de un lugar a otro. Pero la teoría de conmutación, tal y como se la conoce hoy día, se basa en una función económica anticuada y es inoperante para el diseño de los circuitos integrados.

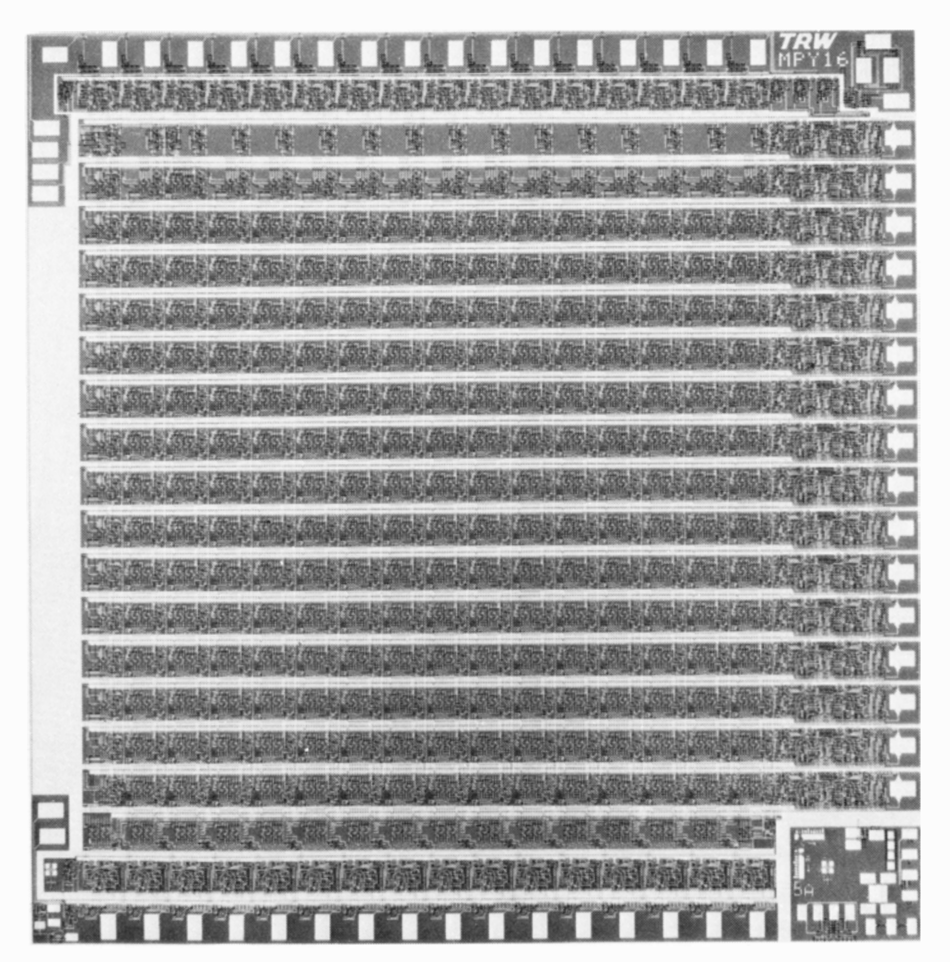
La teoría de la conmutación es todavía menos útil en los diseños que tratan de combinar circuitos integrados dentro de sistemas mayores. En muchos casos, cuesta más probar, empaquetar e inter-



EL COSTE DE UN CIRCUITO INTEGRADO comprende sólo una pequeña parte del coste de un sistema completo. Según vemos arriba, el coste de una matriz típica de circuito integrado vale 10 centavos de dólar. Suponiendo una producción con un 20 % de pastillas buenas, el coste por pastilla, incluyendo empaquetado y pruebas, alcanza 1,60 dólares. Si cada uno de los 20 circuitos impresos incluyera 100 pastillas, el coste unitario por pastilla se duplicaría, al atribuirle la parte proporcional de coste de la placa, del panel del armario y de la fuente de alimentación.



GRUESA MARAÑA DE CABLES, que cubre el panel posterior de un gran calculador de tipo general, en este caso el CRAY-1 de Cray Research, Inc. Trasladar datos por los cables lleva tiempo y vale dinero; el grueso de la maraña representa una dificultad a la hora de las reparaciones. Al reordenar los elementos de un calculador, de forma que los cables queden todos paralelos, se reduce considerablemente la complejidad y el grueso de la maraña.

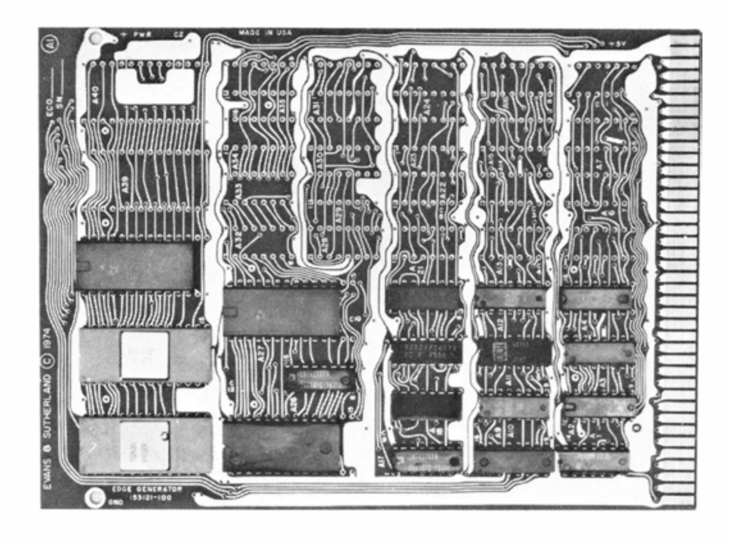


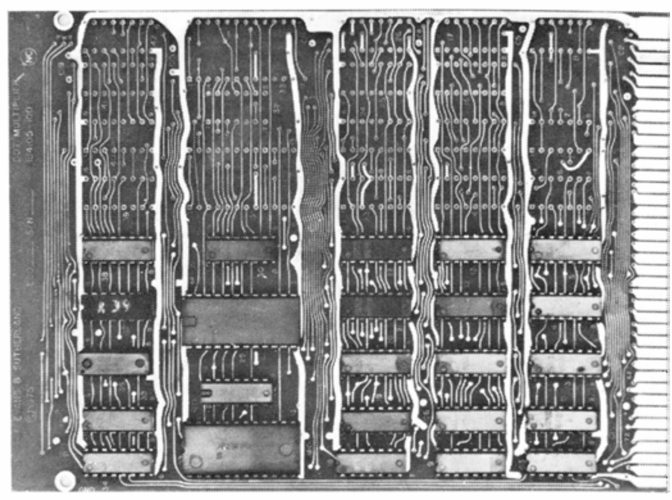
LA REGULARIDAD es una característica de los circuitos de memoria y de ciertos circuitos aritméticos, como este multiplicador matricial de 16 bits, construido por TRW Inc. La pastilla, de unos 7 milímetros, contiene 18.000 transistores y resistencias. Regularidad implica densidad de elementos.

conectar los circuitos integrados que fabricarlos. Estos costes son prácticamente independientes del trabajo particular en el que el circuito está involucrado. Aun prescindiendo del coste del circuito, las conexiones entre las pastillas de un circuito integrado con otro son mucho más lentas que las conexiones en una sola pastilla. Partiendo de un catálogo de circuitos normalizados, se tiende fuertemente a utilizar los circuitos integrados más complejos, porque de esta manera necesitará menos de ellos, con lo que el gran coste de montarlos e interconectarlos quedará reducido. De hecho, los diseñadores especifican circuitos integrados que contienen elementos superfluos, ya que no ven ninguna ventaja económica en suprimirlos. La teoría de la conmutación no tiene nada útil que decir frente a esta problemática de costes y velocidades.

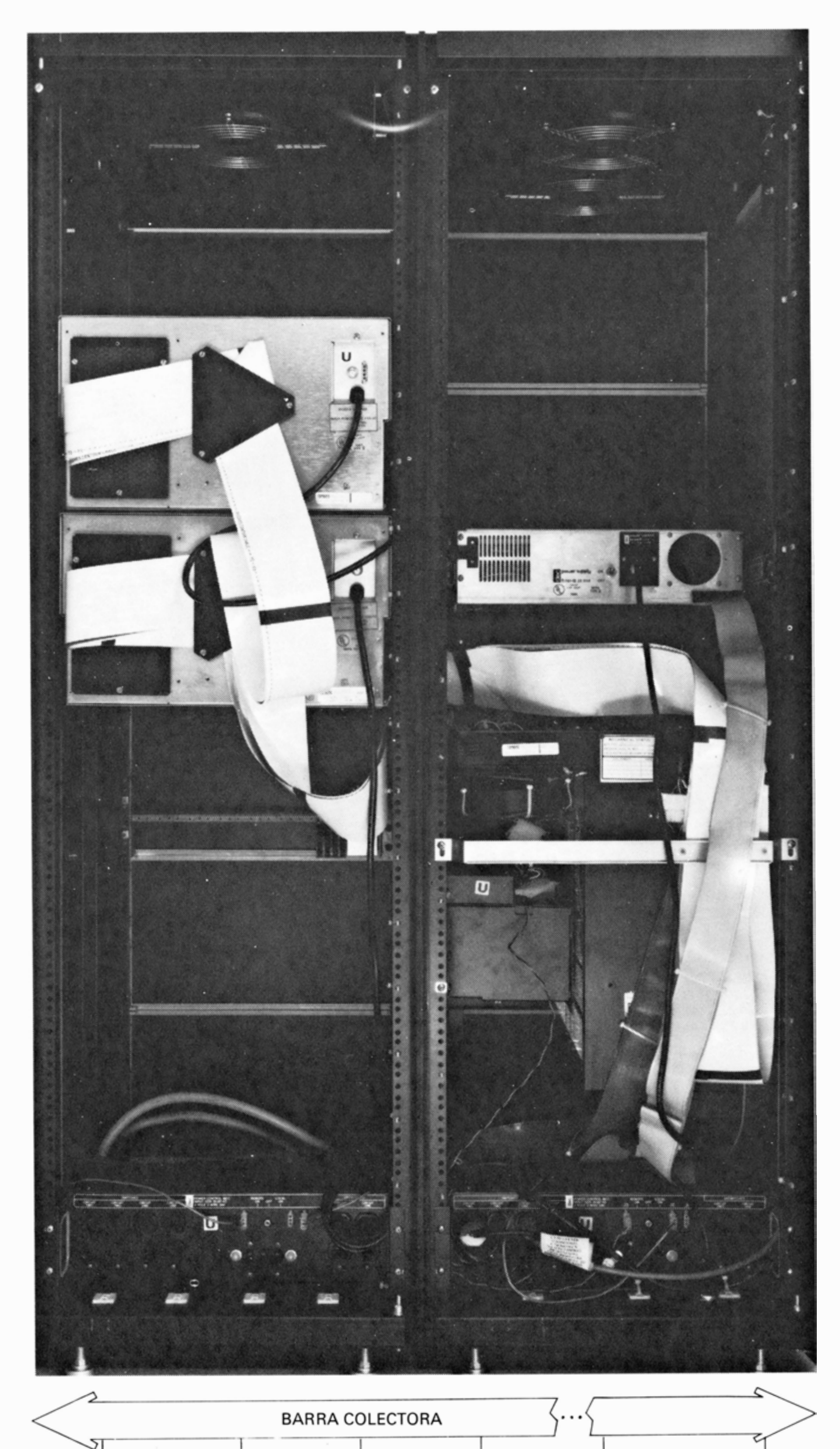
Aunque el coste de las conexiones haya sido ignorado hasta ahora en los resultados teóricos de la ciencia de la informática, debe reconocerse que ha representado cierto papel en la mente de los diseñadores prácticos. Seymour Cray, creador de muchos de los ordenadores más potentes, cita "el grueso de la madeja" y "la eliminación del calor" como los dos mayores problemas en el diseño de las máquinas. Es obvio que someter a control la geometría de las interconexiones es esencial. Si se consigue que las conexiones sigan diseños regulares, se podrán fabricar por métodos menos caros, podrán ocupar menos espacio y ser, en definitiva, más rápidas. Si la geometría de las interconexiones no se planifica cuidadosamente, el espacio que requerirán crecerá cuadráticamente con el número de los elementos lógicos que han de interconectar. Este crecimiento no lineal se debe a que los sistemas mayores requieren más cables, que además, en promedio, serán más largos. Debido a que los canales de interconexión aumentan en número y en longitud, el área o volumen total dedicado a la comunicación acaba siendo desmesuradamente grande: si doblamos el número de puntos a interconectar, distribuidos al azar, hay que cuadruplicar el espacio destinado a las comunicaciones. Para acomodar mayor número de cables en un circuito impreso, generalmente hay que usar placas mayores, las cuales, por lo general, dejan mayor espacio entre los componentes que las menores; una gran ciudad sufre mayor congestión de tráfico que un pueblecito.

No sólo las conexiones largas ocupan un espacio desproporcionado, sino que también funcionan más lentamente que





PLACAS DE CIRCUITO IMPRESO, diseñadas para minimizar la importancia de los caminos de conexión. La regularidad hace disminuir gran cantidad de cableado *(arriba)*. Un circuito de 127 × 177 milímetros de lógica irregular, construido por Evans E. Sutherland Computer Corporation *(arriba, izquierda)*, se compara con una placa de memoria, más regular, del mismo tamaño *(arriba, derecha)*. Cuanto mayor sea la placa, mayor será la importancia del cableado. Un microprocesador LSI-11 de 216 × 254 milímetros de Digital Equipment Corporation registra el área ocupada por las interconexiones, en una placa grande de tipo convencional. Si la densidad del empaquetado de los circuitos de memoria, cableados con regularidad *(arriba, derecha)*, se pudiese haber conseguido a lo largo y ancho de la placa, ésta se habría reducido a la mitad.

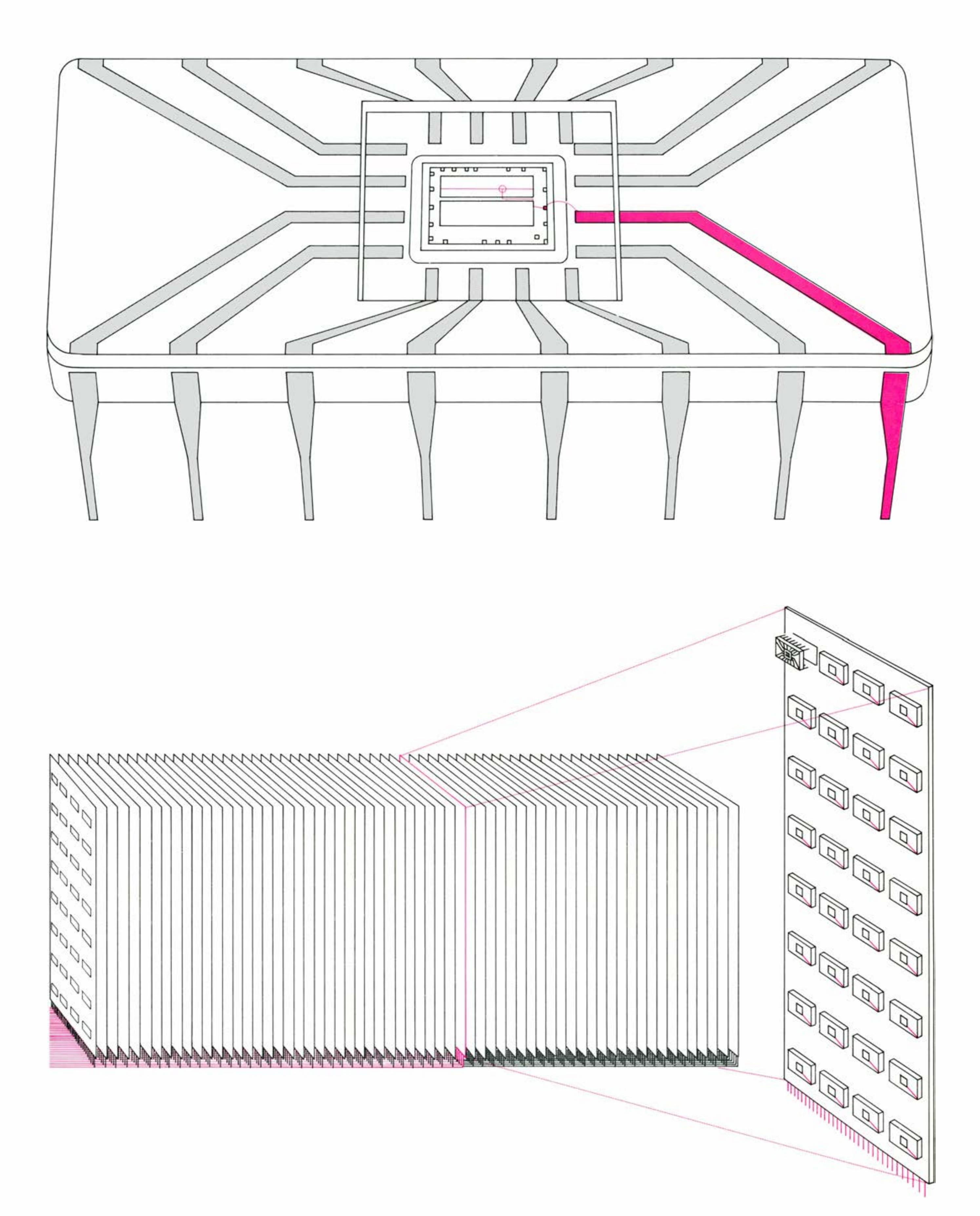


BARRA COLECTORA DE COMUNICACION. Conecta una unidad central de proceso de un ordenador a los módulos de memoria y a otras unidades periféricas *(arriba)*. **Es un tipico cable plano de entre 20 y 100 alambres que se empalman a medida que va pasando por las diferentes unidades. El grabado muestra un UNIBUS, de Digital Equipment Corporation** *(es un cable plano, amplio, de color claro)*, **que conecta dos unidades de memoria de disco** *(arriba-izquierda)* **con la unidad central de proceso** *(abajo, derecha)* **de un calculador PDP-11/40 de DEC.**

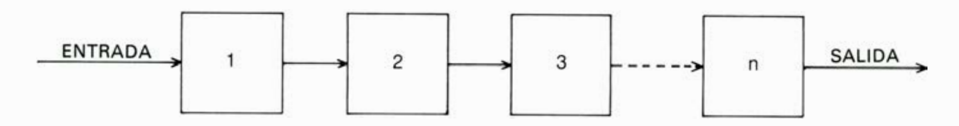
las cortas. Esto se debe a que, incluso a la velocidad de la luz, las señales necesitan cierto tiempo para viajar a lo largo de las conexiones, y por otro lado, a que cuanto más largas son éstas más energía acumulan. (Dentro de los circuitos integrados, el límite de velocidad fijado por la velocidad de la luz no tiene demasiada importancia porque las distancias son cortas comparadas con los tiempos de conmutación de los elementos lógicos; sin embargo, los retrasos debidos a la acumulación de energía sí son importantes.) Antes que una señal pueda conmutar de un estado eléctrico a otro, la energía acumulada en el conductor debe desaparecer y disiparse en forma de calor. Se dispone, entonces, de dos alternativas: o bien se diseña un circuito de envío de la señal más potente, capaz de dar la potencia necesaria para conmutar cables largos a gran velocidad, o bien se aceptan los mayores retrasos que un circuito de envío menos potente provocará al conmutar sistemas mayores de energía. Hay que tener muy presente que los circuitos de envío más potentes no dejan de tener sus inconvenientes. Por un lado, ellos han de ser lanzados a su vez, y, por otro poseen un tamaño mayor y se muestran más lentos que los pequeños.

Pero todavía hay más. El calor generado por los circuitos de envío más potentes tiene que disiparse por alguna estructura, la cual, a su vez, ocupará espacio. Es posible que la energía de las señales requerida en una tecnología determinada y el tamaño de las estructuras necesarias para disipar el calor, establezcan una cota superior a la complejidad de los sistemas que se pueden construir con dicha tecnología. Por encima de esta cota, el incremento de longitud del cable requerido para crear el espacio necesario para introducir el circuito de envío mayor puede exceder al incremento de longitud de cable que se consigue con este mismo circuito. Hasta ahora no hay teoría sobre los límites de velocidad y complejidad que esta posibilidad impone.

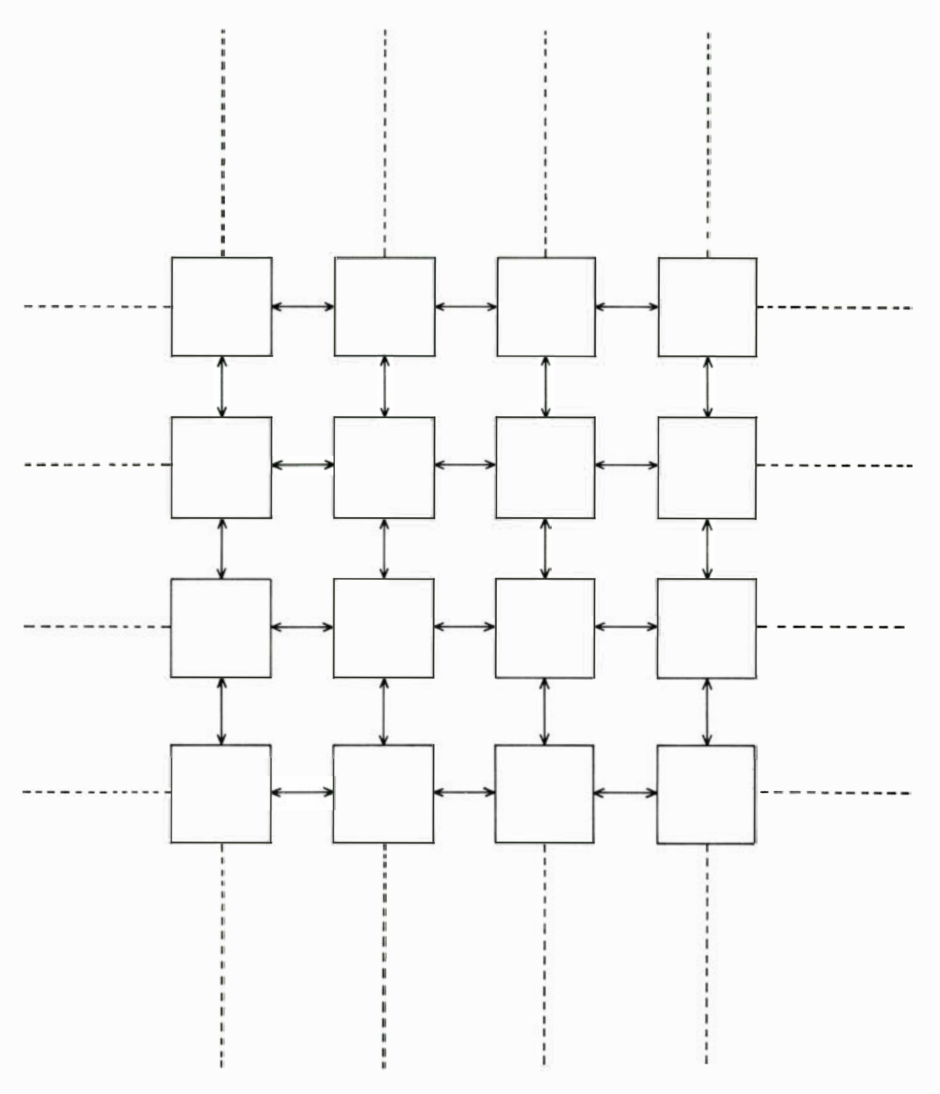
A través de la construcción de esquemas de interconexiones muy regulares puede evitarse el crecimiento desproporcionado de las interconexiones. Hay ya una tendencia hacia esquemas de cableado muy regulares dentro de circuitos integrados y en la interconexión entre circuitos. Las memorias de sólo lectura, por ejemplo, desempeñan funciones de lógica compleja e irregular mediante esquemas muy regulares de circuitos integrados. Regularidad que es buena, no sólo porque simplifica la especificación de



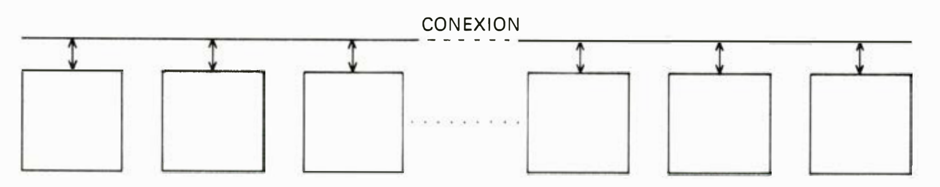
TIPICA PASTILLA DE MEMORIA que tiene 16.384 bits, situados en una matriz de 128 × 128 *(arriba)*. Se accede a una fila completa de 128 bits de golpe, pero un selector permite que solamente pase 1 bit a través del punto de salida *(color oscuro)*. Un sistema de memoria típico se consigue con 2048 pastillas de ésas, es decir, 64 grupos de 32 *(abajo)*. Solamente 32 pastillas pueden situar sus salidas en los 32 cables que unen la barra colectora con el procesador central. De los 262.144 (128 × 2048) bits que se mueven menos de 1 milímetro en cada pastilla, únicamente 2048 avanzan cerca de 3 milímetros para salir de la misma. Unicamente 32 avanzan cerca de un metro para salir del procesador. En otras palabras, la barra colectora sólo utiliza un 8000-avo de la banda pasante disponible de la pastilla de memoria.



PROCESADOR TELESCOPADO, que constituye uno de los tres tipos de procesador paralelo, ilustrados en esta página, que ha resultado ser eficaz. En un procesador telescopado, los datos pasan desde los elementos de un procesador especializado al siguiente, y cada elemento realiza una operación sucesiva de datos. El telescopado se parece a una línea de montaje: todas las operaciones se van sucediendo simultáneamente, aunque sobre distinto material. La configuración en telescopado es la mejor siempre que tenga que realizarse el mismo tipo básico de operaciones.



EL PROCESADOR MATRICIAL muestra su carácter práctico en la realización de muchos e idénticos procesos sobre distintos elementos de datos. Los procesadores reciben las mismas instrucciones, como si se tratara de una compañía de soldados numerándose. La limitación estriba en que los cálculos individuales sólo dependen de los datos propios de cada elemento y de los de sus vecinos. Esta facultad resulta efectiva en ciertos tipos de operaciones, verbigracia, las simulaciones meteorológicas, en las que solamente las interacciones atmosféricas locales son significativas.



PROCESADORES INDEPENDIENTES conectados en un canal de comunicación. Constituyen el sistema de disposición más flexible para la ejecución en paralelo de distintas operaciones. A cada procesador se le asigna una tarea, a medida que se requiera, tal como sucede en una granja con los trabajadores no especializados. El sistema funciona de maravilla cuando cada elemento puede rendir mucho sin necesidad de comunicarse con otros elementos. Los cuellos de botella se presentan cuando las tareas exigen que los elementos trabajen en equipo.

estas funciones, sino también porque puede ser la disposición más eficiente desde el punto de vista de las interconexiones. Creemos que los esquemas regula informática llegará a ser el estudio de la regularidad de tales estructuras.
la informática llegará a ser el estudio de la regularidad de tales estructuras.

La arquitectura de un ordenador típico incluye una única unidad lógica que comunica con una única memoria de acceso aleatorio, a través de entre 20 y 100 cables combinados en una barra colectora o bus, los cuales, como su mismo nombre indica, sirven de transporte público para los datos; aunque, de hecho, una comparación más exacta sería una línea de teléfono de conexión múltiple. La barra colectora de comunicación es un cable flexible de entre 15,24 y 30,48 metros de largo. A la barra colectora se le asigna un protocolo de señales, de modo que todas las unidades a las cuales está conectada comuniquen entre sí colectivamente evitando interferencia de unas con otras. La gran ventaja de una estructura de barra colectora en un ordenador estriba en que cualquier unidad conectada a la barra colectora puede comunicarse directamente con cualquier otra. Más aún, el protocolo de la estructura de barra colectora puede sobrevivir a varias generaciones del desarrollo de las máquinas, de modo que una línea de equipo de cálculo puede incorporar nuevos dispositivos de almacenamiento nuevas unidades de entrada-salida e incluso nuevos elementos de proceso. Además, el número de elementos de conmutación destinados a la comunicación, en cada unidad, con la barra se minimiza porque cada unidad necesita comunicarse sólo con la barra colectora para enviar mensajes en un sentido o en otro.

La estructura de barra colectora tiene el inconveniente de que ofrece una comunicación en cuello de botella. Consideremos un típico calculador que tenga, digamos, un millón de palabras (32 millones de bits) de memoria de circuitos integrados; construida por medio de 2048 circuitos, cada uno de los cuales albergue 16.000 bits. Una referencia a la memoria puede consultar potencialmente los valores de 128 bits de cada uno de los 2048 circuitos integrados que constituye la memoria. De este cuarto de millón de bits a los cuales se accede en el circuito integrado, sólo 2048 (uno de cada circuito integrado) sale fuera del paquete de circuitos integrados, y de estos 2048, sólo 32 (o sea una palabra) desembocan sobre la barra colectora de comunicación y, a través de ésta, sobre la unidad central de proceso. Se supone que la barra

colectora de comunicación conecta la memoria con la unidad lógica de proceso; como acabamos de ver, más que unirlos los separa. Cada acceso de memoria de un ordenador grande desaprovecha accesos a millares de bits cuando escoge unos pocos para enviar a través de la barra colectora de memoria a la unidad central de proceso. Este despilfarro se tolera por dos razones. Primero, porque simplifica nuestro concepto de la máquina y se adapta a nuestra inclinación natural de hacer las cosas una a una. Segundo, nos suministra una sola y simple interconexión entre las distintas partes de la máquina.

Pagamos un elevado precio por estas ventajas. En la época en que las memorias y los elementos lógicos de proceso se hacían por medio de tecnologías distintas teníamos pocas alternativas. Hoy día, sin embargo, en que los circuitos integrados de silicio dominan tanto en las memorias como en los elementos lógicos de proceso de los ordenadores, no hay justificación para continuar aceptando este despilfarro. Se puede distribuir la barra colectora de memoria entre varios miles de circuitos integrados, con lo que se consigue, de forma efectiva, dar a cada elemento lógico la memoria que necesite; así pues, mueve la información menos de un milímetro desde la memoria hasta los dispositivos de proceso, localizados conjuntamente en los mismos circuitos integrados en cada una de los miles de pastillas. Estamos ahora empezando a explorar sistemas con este tipo de arquitectura no convencional. Para emplearlos de forma eficaz, debemos aprender a compaginar la complejidad de los problemas a resolver con los simplicísimos esquemas de conexión posibles en los sistemas que podemos construir.

Se denominan procesadores en paralelo aquellas máquinas en las que un gran número de elementos lógicos trabajan simultáneamente. (En cierto modo, cualquier máquina de cálculo es un procesador en paralelo. Los bits independientes. que conjuntamente representan un número, se mueven simultáneamente a lo largo de las interconexiones en paralelo; una suma binaria se realiza en un circuito sumador que opera con todos los bits de un número a la vez; la multiplicación se realiza por sumadores secuenciales o, en modelos más rápidos, mediante inclusión de varios circuitos de distintos sumadores, los cuales trabajan entre sí en paralelo. Niveles de paralelismo más avanzados que estos básicos son todavía raros en los ordenadores actuales.) El tipo más simple de ordenador paralelo que hoy en día disponemos tiene unos cuantos procesadores independientes, que operan sobre una única memoria; un típico gran sistema de cálculo tiene entre dos y varias docenas de procesadores trabajando simultáneamente. Sin embargo, estos procesadores sirven para funciones prácticamente independientes, y su misma existencia puede quedar oculta al usuario. Por ejemplo, un procesador puede dedicarse a las comunicaciones con la consola del operador, para manipular las unidades de entrada-salida de cintas magnéticas y discos magnéticos o la planificación de recursos del procesador central. Estos sistemas de "multiproceso" ejercen escasa incidencia en los algoritmos del usuario.

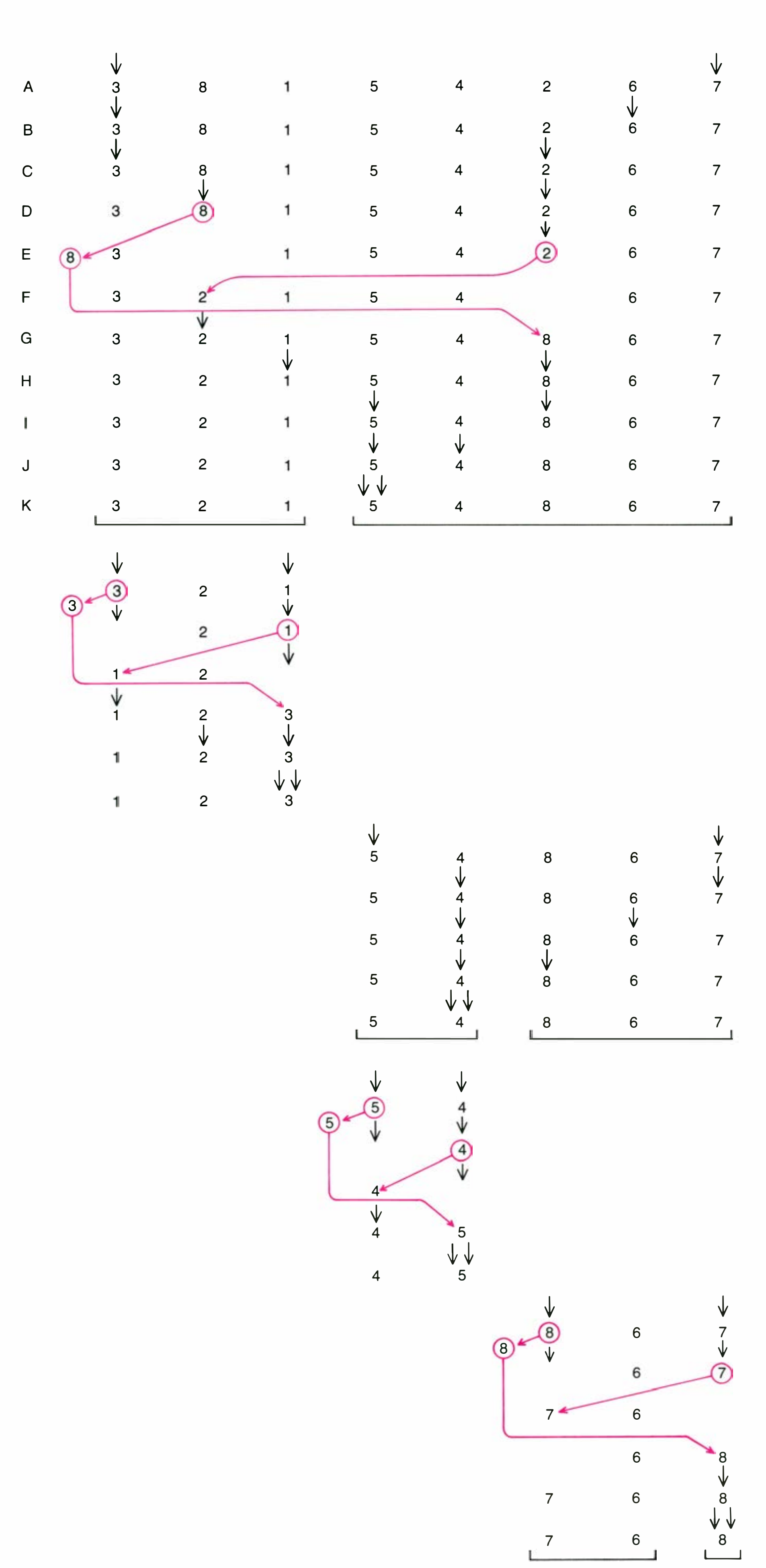
Se han construido tres clases de sistemas que pueden considerarse como procesadores paralelos. En uno de ellos, el denominado procesador telescopado ("pipeline"), diversos elementos de proceso, cada uno de los cuales está especializado en una tarea determinada, se hallan conectados en secuencia. El trabajo a realizar fluye a través de los procesadores, de la misma manera que las piezas avanzan a lo largo de una cadena de montaje. La comunicación es sencilla porque el flujo de información pasa a lo largo de caminos fijos y media muy poca distancia entre los procesadores. Un procesador telescopado consigue eficiencia por la misma razón que lo hace una cadena de montaje: las funciones son especializadas y las comunicaciones mínimas. El telescopado permite a las secciones aritméticas de los ordenadores muy rápidos procesar secuencias de números con una velocidad muy alta. Los procesadores telescopados pierden en eficacia cuando los trabajos a realizar son altamente variables.

En una segunda forma de procesadores paralelos, muchos elementos de proceso idénticos se montan para trabajar en distintas partes de un problema, bajo el control de una sola secuencia de instrucciones. Se han construido distintos ordenadores de este tipo, de los cuales el más conocido es el ILLIAC IV. En una reciente reunión de la Rand Corporation sobre simulaciones hidrodinámicas acaba de proponerse un procesador moderno de este tipo. Dicha máquina hipotética incluiría 10.000 procesadores, con capacidad aritmética y de memoria individual; cada uno estaría construido en un único circuito integrado, y, todos, bajo el mando de un dispositivo de instrucción común. Todos los procesadores ejecutarían las órdenes a un compás rígido. Los procesadores se montarían en una matriz cuadrada de 100 × 100; cada uno enviaría datos solamente a sus vecinos, al norte, sur, este

y oeste en la matriz; y ello según el modo serial, lento, con un único cable en cada dirección. A nuestro entender, esa máquina necesitaría del orden de cinco microsegundos para comunicar un sencillo número de 64 bits desde un procesador a su vecino, lo cual resultaría francamente lento para las velocidades actuales. Claro que se podrían comunicar 10.000 de tales números en un período de cinco microsegundos. Tardaría también otros cinco microsegundos en realizar una multiplicación, pero de nuevo podrían realizar 10.000 multiplicaciones simultáneas con un promedio de dos mil millones de multiplicaciones por segundo. Este tipo de máquinas se llaman procesadores matriciales o procesadores de flujo único de instrucciones y múltiple flujo de datos. Se adaptan muy bien para tareas de tipo muy regular, verbigracia: cálculos hidrodinámicos, simulaciones numéricas meteorológicas e inversión de grandes matrices.

Un tercer tipo de procesador paralelo lo constituyen procesadores separados, independientes, bajo estructuras de control separadas y autónomas que realizan subtareas independientes, comunicándose datos e instrucciones según convenga. La aparición de los microprocesadores ha sugerido a mucha gente la posibilidad de idear sistemas que consistan en miles de microprocesadores independientes, que operen coordinadamente en grandes trabajos. Se han construido algunos de estos procesadores de múltiple flujo de instrucciones y múltiple flujo de datos, pero sus propiedades apenas se conocen.

El estímulo para diseñar y usar procesadores paralelos de cualquiera de los tres tipos estriba en descubrir cómo se adaptan los esquemas simples de comu-

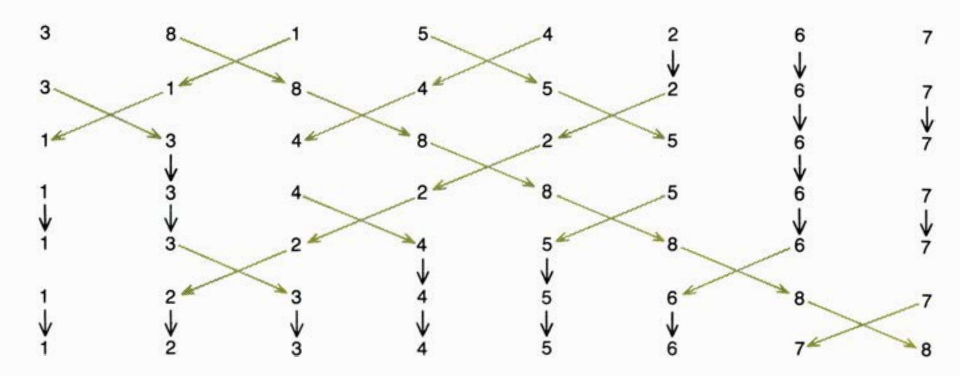


EL "QUICKSORT" es uno de los típicos algoritmos secuenciales para ordenar números en progresión ascendente. Se comparan los números apuntados por las flechas con el número situado más a la izquierda. Si la primera flecha por la izquierda supera el valor de referencia *(fila A)*, se le desplaza hacia la izquierda, hasta encontrar un valor inferior a la misma *(C)*. Luego la flecha izquierda se desplaza hacia la derecha hasta encontrar un valor superior a la referencia *(D)*. En esta circunstancia se intercambian los números apuntados *(E-G)*. Se continúa el proceso hasta que las flechas descansan sobre el mismo número *(K)*. En este estadio, todos los numeros a la derecha de las flechas son mayores que la referencia, y los que están a la izquierda, menores o iguales. El mismo algoritmo se aplica a cada subconjunto; se muestran los cinco pasos más que hay que aplicar para alcanzar la ordenación total. Ordenar un número requiere $n(\log_2 n)$ comparaciones entre números que pueden estar situados en posiciones muy distanciadas. El coste dominante de la ejecución del algoritmo es el coste de las comunicaciones.

nicación en el interior del procesador a las tareas de comunicación inherentes al problema que se quiera resolver. A medida que la tecnología de circuitos integrados progrese, habrá circuitos individuales de velocidad y complejidad crecientes. Sin embargo, no se tiene ninguna mejora en perspectiva para las pérdidas por retrasos inherentes a la comunicación de información entre circuitos. Conseguir mejores comunicaciones requeriría mayor número de conexiones entre los circuitos integrados, más espacio para alojarlas y más potentes y largos circuitos de envío de comunicación, y como consecuencia, mayor disipación del calor. Para conseguir las máximas prestaciones de un sistema de cálculo complicado, los programadores tendrán que enfrentarse con las limitaciones de comunicación que quedan impuestas por la realidad física. No se puede conseguir una comunicación de elevado rendimiento entre elementos cualesquiera; el programador habrá de adaptar la formulación de su problema a las interconexiones disponibles. Aunque es un trabajo difícil, conseguirlo dará como resultado potencias de proceso nunca logradas.

Creemos que así como una importante parte de la informática de hoy día tiene que ver con la consecuencia de instrucciones distribuidas cronológicamente, de la misma forma un importante aspecto de la informática del futuro será el estudio de conjuntos de comunicaciones distribuidas en el espacio. Si los procesadores sólo pueden dialogar con sus vecinos inmediatos, ¿qué clase de disposiciones son posibles? Obviamente los procesadores se pueden cablear en una cadena o sucesión lineal. Estos procesadores pueden operar en la forma telescopada descrita más arriba, pasando cada uno la información a su vecino, o pueden realizar operaciones comunes bajo la orden de un dispositivo central. Las conexiones con el vecino son eficaces para ciertas tareas, como la ordenación, en la cual son suficientes las comunicaciones de tipo local. Como otra alternativa, se pueden conectar los procesadores en matriz, de forma que cada procesador tenga más de dos vecinos. Estas matrices poseen una estructura básica análoga a la cristalina, y son posibles varias formas de comunicación local. En nuestro laboratorio del Instituto de Tecnología de California estamos considerando las posibilidades de distintos sistemas de comunicación que se podían incorporar a diferentes estructuras.

Hace varios años, R. S. Gaines y C. Y. Lee, que trabajaban en los Bell Laboratories, describieron tres tipos de caminos de interconexión. Un tipo de interconexión tiene enlaces que son comunes a todos los procesadores: es eficaz para enviar órdenes a todos los procesadores y radiar ciertos valores que pueden serles importantes en el curso de los cálculos. Un segundo tipo de comunicación permite a cada procesador dialogar simultáneamente con sus vecinos. Un enlace de este tipo puede manipular la información requerida en el telescopado; si cada procesador tiene su propia función de memoria se puede abrir una brecha en cualquier lugar de la memoria, trasladando información simultánea desde la localización de dicha brecha a cualquier otra parte. Un tercer tipo de enlace de comunicación permite a los elementos del proceso hablar colectivamente acerca de sus propios resultados. Este enlace puede decir si ningún procesador, un procesador, o más de uno, cumple una determinada condición, cuál de los procesadores contiene el valor más pequeño y cuáles están en el principio o en el final de una determinada cadena.



LA EJECUCION EN PARALELO acelera los procesos de ordenación. En un algoritmo de este tipo, se comparan entre sí parejas de números adyacentes y se intercambian si el de la izquierda es mayor que el de la derecha. (En la primera fila, una pareja queda definida de modo que el número de la izquierda esté en una columna par; en la segunda fila, el miembro de la izquierda de cada pareja está en una columna impar, y así sucesiva y alternadamente.) Ordenar todo el conjunto requiere $n^2/2$ comparaciones, siempre entre elementos vecinos. Este algoritmo de ordenación por intercambio se caracteriza por su lentitud; pero si las comparaciones e intercambio de elementos se hacen por circuitos integrados unidos a cada elemento de memoria, se pueden hacer N comparaciones de una sola vez, en un único ciclo de memoria, y la ordenación, por consiguiente, quedará realizada en n ciclos como máximo.

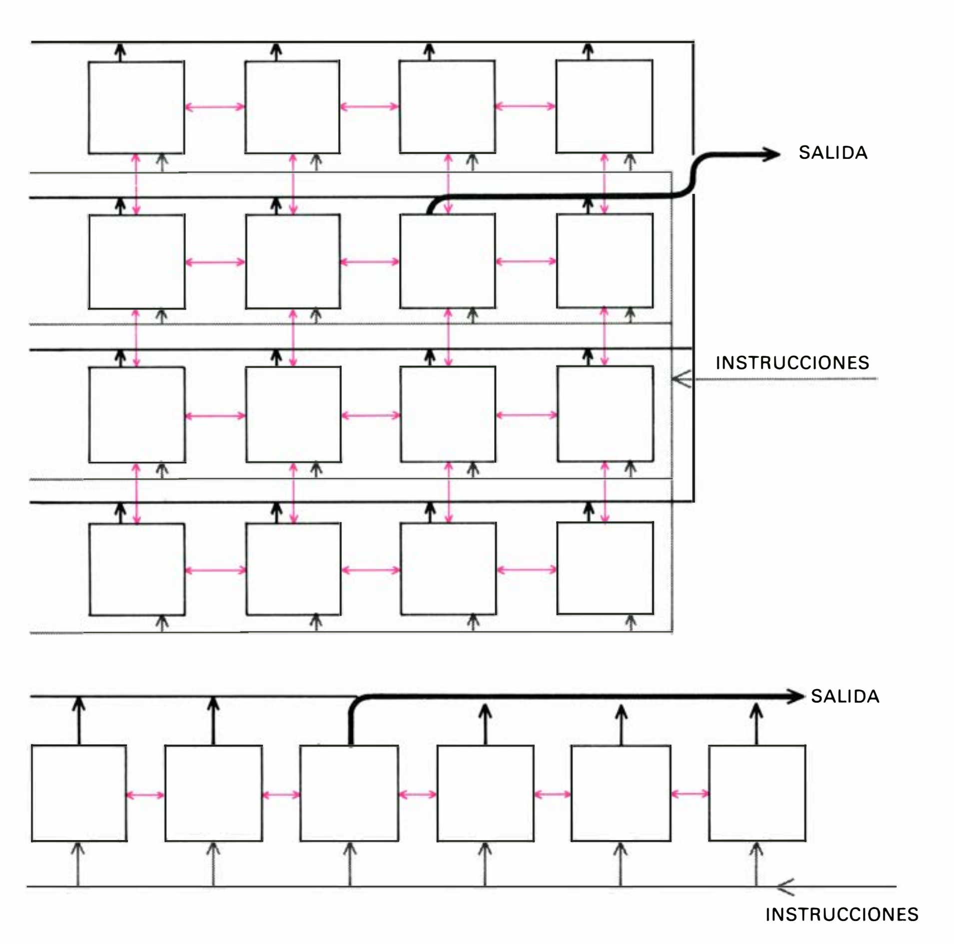
En nuestro laboratorio nos hemos impuesto la tarea de construir y usar algunos de estos procesadores simples en paralelo, dotados de esquemas de interconexión mono, bi y tridimensional.

Esperamos comprender mejor la relación que existe entre las vías de comunicación y el rendimiento de tales procesadores. Nos hemos convencido de que los rendimientos de los procesadores en paralelo pueden depender, de forma crítica, del diseño del esquema de comunicaciones que permite a los elementos del proceso hacer afirmaciones colectivas sobre sus propias acciones. Sin estos enlaces, ¿cómo se puede conseguir encontrar el mínimo valor almacenado en la matriz? ¿cómo se puede identificar el conjunto de procesadores con determinadas propiedades que están situados entre dos procesadores? ¿cómo se pueden conseguir respuestas de procesadores en secuencia, cuando más de uno de ellos tiene algo que decir? Tales enlaces son eléctricamente complejos. O bien involucran a cada procesador local como elemento activo de una tarea global de comunicación, o bien requieren circuitería especializada, que se dedique a recoger tal información; circuitos intermediarios que, inevitablemente, introducen pérdidas de tiempo y coste. Las mejores estructuras para este tipo de comunicaciones parecen ser similares a los circuitos de arrastre de los sumadores paralelos rápidos, pero los costes de comunicación de tales circuitos todavía no se han analizado adecuadamente.

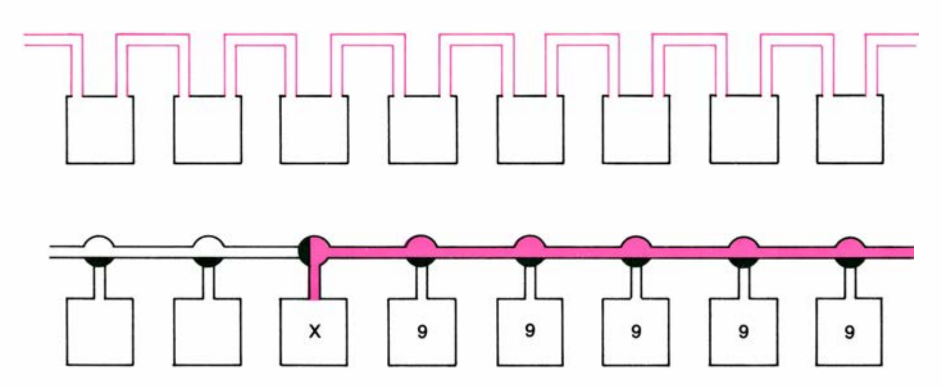
Piedra miliar de la ciencia informática de hoy día es el análisis teórico de los algoritmos secuenciales. Existe un cuerpo de teoría, grande y creciente, sobre cómo seleccionar algoritmos eficientes para máquinas secuenciales. Esta teoría, como era de esperar, concentra su atención en los algoritmos que minimizan el número de operaciones lógicas para realizar una tarea. Por ejemplo, se ha demostrado que, para ordenar números en secuencia, los algoritmos llamados "quicksort" (de ordenación de elementos "muy rápida") son los mejores, porque requieren menor número de comparaciones, concretamente $n\ (\log_2 n)$. Esta teoría supone que todos los elementos almacenados son igualmente accesibles y que el traslado de los datos no cuesta nada.

La verdad es que los datos de la memoria nunca son igualmente accesibles, aunque en un dispositivo de acceso alea-

torio se podrían hacer arbitrariamente asimismo inaccesibles, obligando a que el tiempo de acceso para todos los elementos fuera tan lento como el requerido para los elementos menos asequibles. Debido a que la información, en un dispositivo de memoria aleatorio, debe desplazarse por grandes distancias, la velocidad de flujo de los datos es, en estos casos, inevitablemente más lenta que la que se puede conseguir con un mecanismo de acceso secuencial ordenado. Es más, trasladar datos desde una celda de memoria hasta un circuito de comparación nunca es gratuito; en la mayoría de las máquinas, el tiempo de transporte es muy superior al de comparación. Así por ejemplo, un análisis que intente demostrar que un determinado algoritmo de ordenación es el óptimo, puede consistir en ordenar la tarea de realizar este algoritmo a una máquina subóptima; en una estructura distinta, la ordenación se puede realizar mucho más rápidamente.



LOS CABLES QUE CONECTAN LOS MODULOS de un programador matricial se pueden disponer de tres maneras distintas: Un tipo de conexión difunde información desde un centro de control a todos los módulos *(gris)*. Un segundo tipo *(negro)* traslada la información desde un determinado módulo al control, módulo tras módulo *(flecha gruesa en negro)*. Un tercer tipo envía datos de cada módulo a su vecino *(color)*; aquí, todos los módulos "hablan" al mismo tiempo.

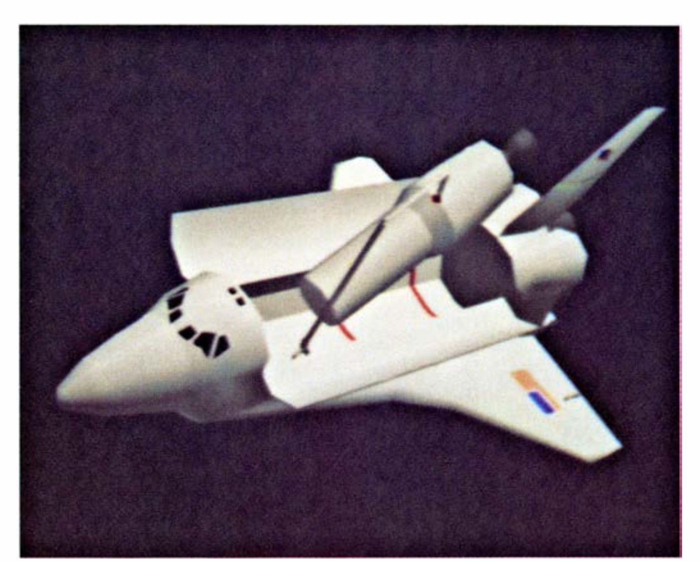


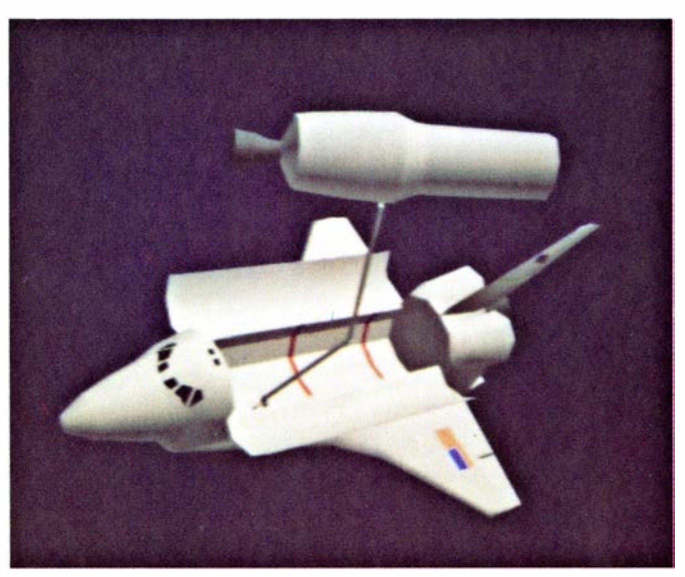
DOS SUBTIPOS del tercer tipo de cableado, que son de uso común. En uno de ellos *(arriba)*, la información atraviesa el módulo en un paso. Se procesa y pasa al siguiente módulo en el paso siguiente. El otro subtipo *(abajo)*, ha sido diseñado sólo para "detectar", y con rapidez, algo relacionado con un módulo (o relacionado con varios módulos colectivamente); la información pasa a través de un módulo durante un único paso de proceso, sin tener que sufrir cambios, a menos que el módulo se encuentre en un estado específico. Este tipo de cableado puede servir para descubrir, por ejemplo, dónde hay nueves en un sumador paralelo. (Ilustración de Alan D. Iselin.)

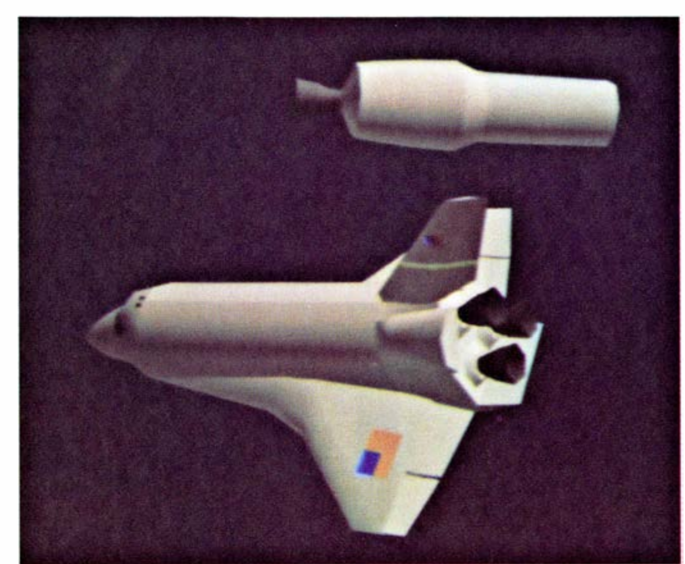
En la teoría de algoritmos, el esfuerzo no se ha orientado todavía hacia cuál sea la verdadera relación entre los costes de computación y los costes de comunicación. Supongamos que se empiece con una placa virgen de silicio, y que se tenga la libertad de colocar cables, puertas lógicas, etcétera, dondequiera. ¿Qué elecciones deberían hacerse para conseguir que una determinada tarea de computación se realice en un tiempo mínimo o en una cantidad mínima de la superficie de la placa? Carecemos en absoluto de base teórica para escoger correctamente de acuerdo con la estructura de cálculo que podría construirse. Tenemos gran caudal de experiencia en una determinada estructura: las máquinas secuenciales con memoria de acceso aleatorio. Puede ser que tales máquinas sean eficaces porque se adaptan, de una forma global, a gran variedad de tareas. Pero se hace cada día más evidente que una estructura en paralelo superaría a los calculadores típicos, en varios órdenes de magnitud, en trabajos adaptados a su ejecución en paralelo. Cuando tales estructuras no tradicionales hayan aparecido se descubrirá un amplio espectro de tareas que se puedan adaptar a estas máquinas. "A priori", uno esperaría que se consiguiesen mejores rendimientos si se prestase más atención a los costes reales del sistema, lo que es decir a los costes de comunicación, en vez de considerar simplemente los costes –en vías de extinción– del proceso lógico.

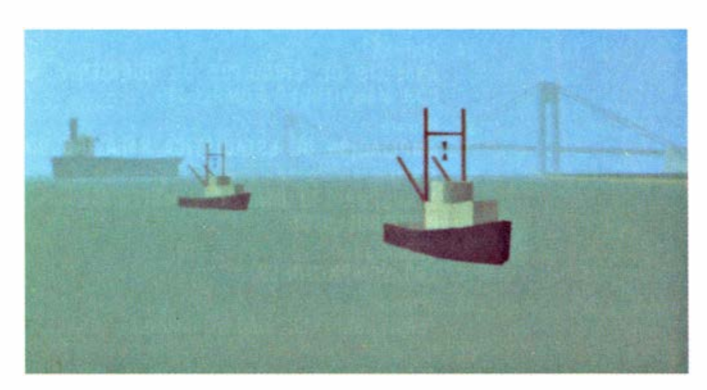
Pensamos que una importante guía para el diseño de las máquinas que han resultado posibles después de la revolución de los circuitos integrados serán las teorías que tengan en cuenta, adecuadamente, los costes de comunicación. Creemos que estas teorías hallarán su base en el estudio de la regularidad, con lo que la informática incluirá un capítulo de teoría análogo al de la topología o la cristalografía. Aunque un desarrollo de este tipo resulta revolucionario en muchos aspectos, en el fondo se trata de la continuación de la búsqueda de regularidad que se venía realizando en todas las tareas de programación. Los científicos del cálculo añadirán simplemente la regularidad geométrica a la regularidad lógica que siempre han conocido y valorado.

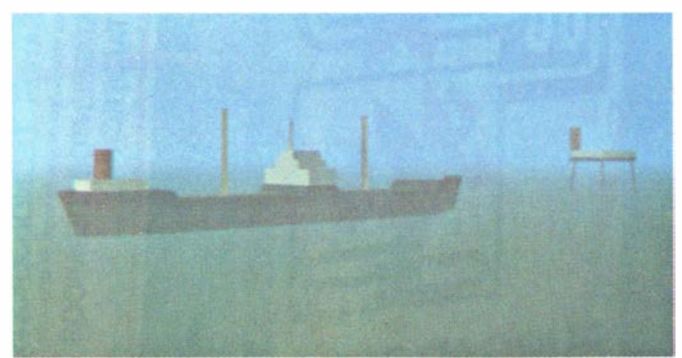


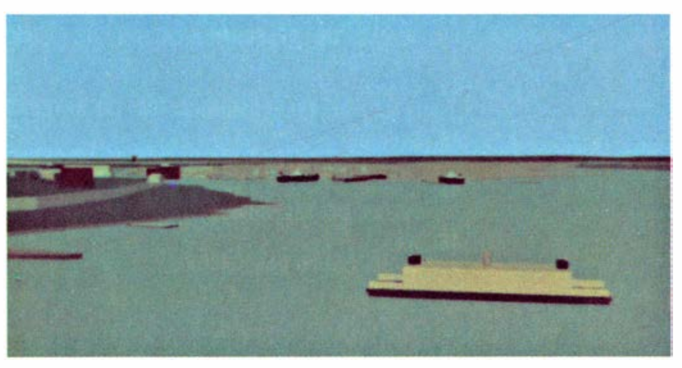


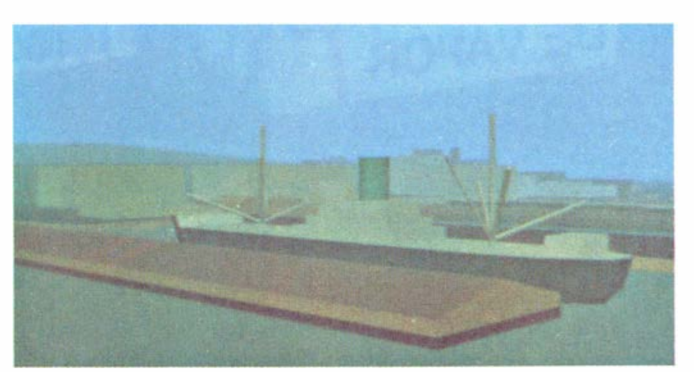












Microelectrónica y ordenador de uso personal

El ritmo de progreso en la microelectrónica hace pensar que dentro de una década mucha gente poseerá un ordenador del tamaño de un cuaderno y la capacidad de un gran ordenador actual. ¿En qué podrá ayudarles ese sistema?

Alan C. Kay

El futuro incremento en capacidad y abaratamiento del coste de los dispositivos microelectrónicos no sólo dará lugar a un soporte físico potente y compacto, sino que comportará también cambios cualitativos en la manera de relacionarse el hombre con el ordenador. En la década de los años 80, adultos y niños podrán disponer de un ordenador del tamaño aproximado de un cuaderno, cuya potencia les permitirá manejar toda la información relativa a sus necesidades. El cálculo y la capacidad de almacenamiento superará con mucho a las virtualidades de los microordenadores actuales: decenas de millones de operaciones básicas por segundo manejarán el equivalente a varios miles de páginas de información impresa.

El ordenador de uso personal puede considerarse como el ejemplo más reciente entre los medios de comunicación humana. Desde que los seres humanos empezaron a hablar han existido distintos medios de almacenamiento, recuperación y manipulación de información. Los medios externos sirven para encerrar los pensamientos internos y comunicarlos, y, a través de procesos de realimentación, construir el camino que sigue el pensamiento. Aunque los ordenadores digitales se diseñaron para realizar operaciones aritméticas, su capacidad para simular los comportamientos de cualquier modelo descriptivo significa que el ordenador, en cuanto medio asimismo, puede simular cualquier otro medio, con tal que los métodos de simulación estén suficientemente bien descritos. Más todavía, a diferencia de los medios convencionales que son pasivos (ni los signos escritos sobre un papel, ni las pinceladas realizadas sobre un lienzo, ni las imágenes de la televisión cambian en respuesta a los deseos del observador), el ordenador se muestra activo: puede contestar las solicitudes y ensayos e incluso puede sostener un diálogo con el usuario.

La evolución de los ordenadores de uso personal ha seguido un proceso similar al de la letra impresa, con la salvedad de que sólo posee 40 años de vida y no 600. Como ocurrió con los manuscritos de la Edad Media, los pesados ordenadores construidos en las dos décadas anteriores a los años 60 fueron escasos, caros y al alcance privilegiado de unos pocos. Así como la invención de la imprenta desembocó en el uso comunitario de los libros reunidos en la biblioteca, la introducción del tiempo compartido en los ordenadores durante los años 60 extendió la capacidad de éstos, que eran caros, disminuyendo su coste de utilización y permitiendo un empleo comunitario. Y al igual que la revolución industrial hizo posible el libro de uso personal, gracias al bajo coste del papel y a la mecanización de la impresión y la encuadernación, es de prever que la revolución microelectrónica de nuestra década dé paso a los ordenadores de uso personal de los años 80, con suficiente almacenamiento y velocidad como para soportar lenguajes de alto nivel y monitores gráficos interactivos.

En principio, los ordenadores de uso personal podrán diseñarse de suerte que personas de todas las edades y profesiones puedan amoldar y canalizar su potencia según sus propias necesidades. Dentro del marco de la hipótesis, el arquitecto podrá simular un espacio tridimensional para reflejar y modificar sus proyectos; el médico podrá almacenar y organizar gran cantidad de información sobre sus pacientes, capacitándole para descubrir importantes relaciones, que, de otro modo, serían imperceptibles; el músico podría escuchar su composición mientras la está creando, sobre todo si le es demasiado difícil interpretarla; el hombre de negocios podrá disponer de un portafolios activo que contenga una simulación del funcionamiento de su compañía; el educador podrá realizar su propia versión de un diálogo socrático con simulación dinámica y animación gráfica; el ama de casa podrá almacenar y manipular archivos, cuentas, presupuestos, recetas y notas; por último, el niño podrá tener una activa herramienta de aprendizaje que le permita acceder fácilmente a un volumen de conocimientos inalcanzables con el mero auxilio de los libros.

¿Cómo sacar partido a la comunicación vía ordenador para hacer frente a las diversas necesidades de los individuos? Para que el ordenador sea verdaderamente "personal", los usuarios han de poder utilizarlo sin tener que recurrir a los servicios de un experto. Las tareas sencillas han de ser sencillas, y las

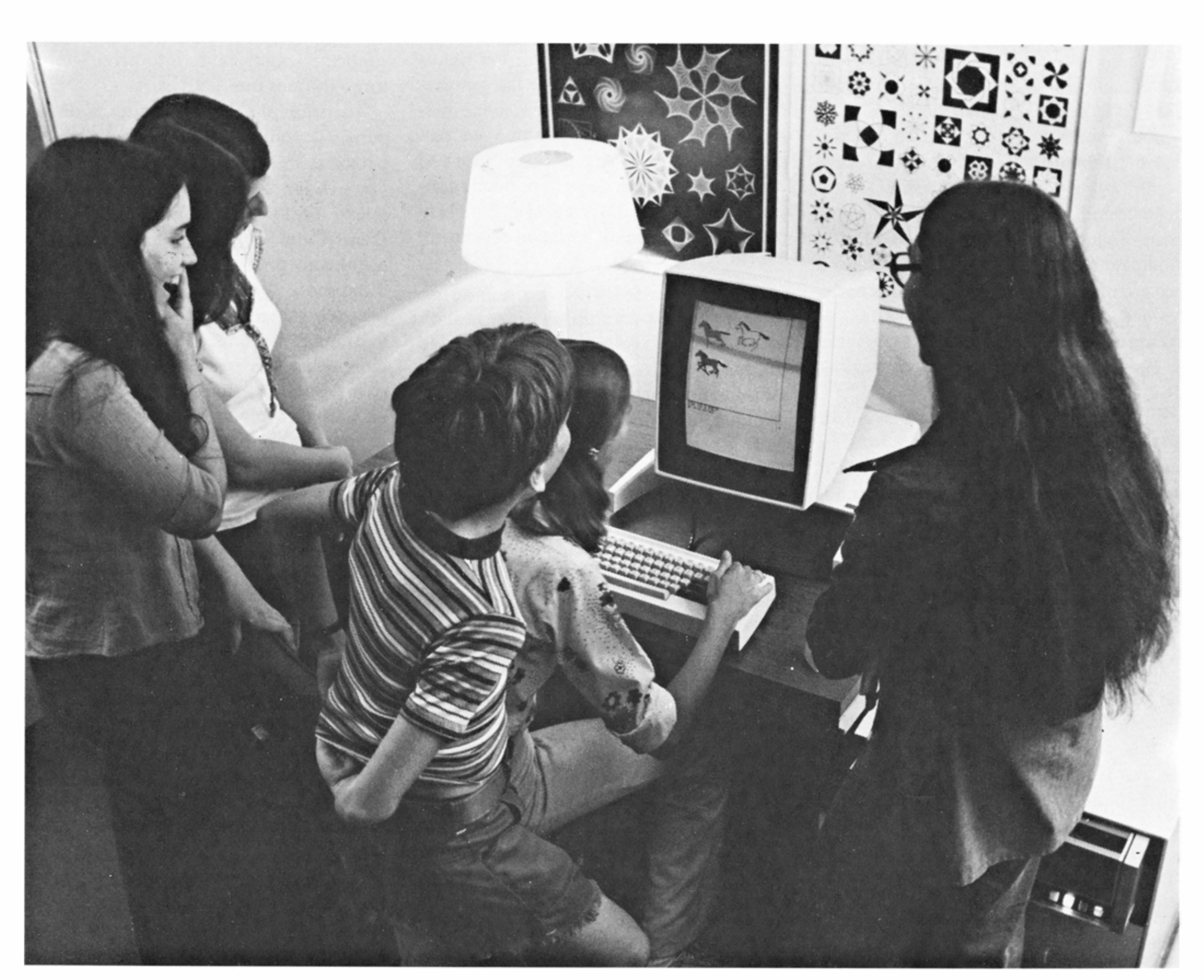
SIMULACIONES DE UN ORDENADOR generadas sobre un monitor de televisión de alta resolución en la Evans & Sutherland Computer Corporation. La calidad de las imágenes, que con el tiempo debieran ser posibles de obtener sobre un ordenador compacto de uso personal, es espléndida. Los dibujos están construidos a partir de dos programas de simulación dinámica que modifican una imagen 30 veces por segundo, para representar el movimiento continuo de objetos en el espacio tridimensional proyectado. La secuencia superior, hecha para la National Aeronautics and Space Administration, muestra un laboratorio espacial despegando del interior del módulo lanzador. La secuencia inferior, hecha para la U.S. Maritime Administration, representa el movimiento de un petrolero en el puerto de Nueva York. La capacidad de un ordenador de uso personal para simular un fenómeno real o imaginado, hará de él un nuevo medio de comunicación.

complejas, posibles. Aunque un ordenador de uso personal vendrá provisto de unas simulaciones ya creadas, como si se tratara de una enciclopedia general, el amplio abanico de edades y conocimientos de los usuarios en potencia hará difícil cualquier anticipación directa de sus necesidades. De este modo, el problema central de los cálculos personales es que los no expertos, sin lugar a dudas, tendrán que hacer algo de programación, si quieren que su ordenador les sirva para algo más que para echarles una mano en un caso esporádico.

Para comprender mejor los problemas y ventajas potenciales de los ordenadores de uso personal, diseñamos en el Xerox Palo Alto Research Center, un sistema experimental para cálculos personales. Allí construimos bastantes sistemas de éstos y estudiamos el uso que niños y adultos hacían de ellos. El soporte físico tenía la misma capacidad que el ordenador tamaño cuaderno que habíamos previsto para los años 80, aunque forzosamente mayor. Para el soporte lógico se buscó un nuevo sistema de lenguaje interactivo que dio en llamarse SMALLTALK.

En el diseño de nuestro sistema de cálculo personal, recogimos las investigaciones hechas en las postrimerías de la década de los 60. Por aquellas fechas me hallaba trabajando, con Edward Cheadle, en la Universidad de Utah, en donde diseñamos el FLEX, que fue el primer ordenador de uso personal que soportaba directamente lenguajes orientados a gráfica y simulación. Aunque el diseño de FLEX nos gratificó mucho, no abarcaba lo suficiente como para servir a una amplia variedad de usuarios no expertos. Empezamos luego a interesarnos por los esfuerzos que estaban llevando a cabo Seymour A. Papert, Wallace Feurzeig y otros en el Instituto de Tecnología de Massachusetts y en Bolt, Beranek and Newman, Inc.; andaban con el desarrollo de un medio de aprendizaje basado en ordenadores con el cual los niños, al tiempo que aprendieran, pudieran encontrar satisfacción y fuente de diversión. Trabajando con un gran ordenador de tiempo compartido, Papert y Feurzeig idearon un lenguaje de ordenador sencillo y potente llamado LOGO. Con este lenguaje, los niños (de edades comprendidas entre los 8 y los 12 años) podían escribir programas para controlar un sencillo generador de música, un robot-tortuga que podía arrastrarse por el suelo y dibujar rayas, y, por último una imagen en televisión de la tortuga, que podía hacer las mismas cosas.

Después de observar este proyecto,



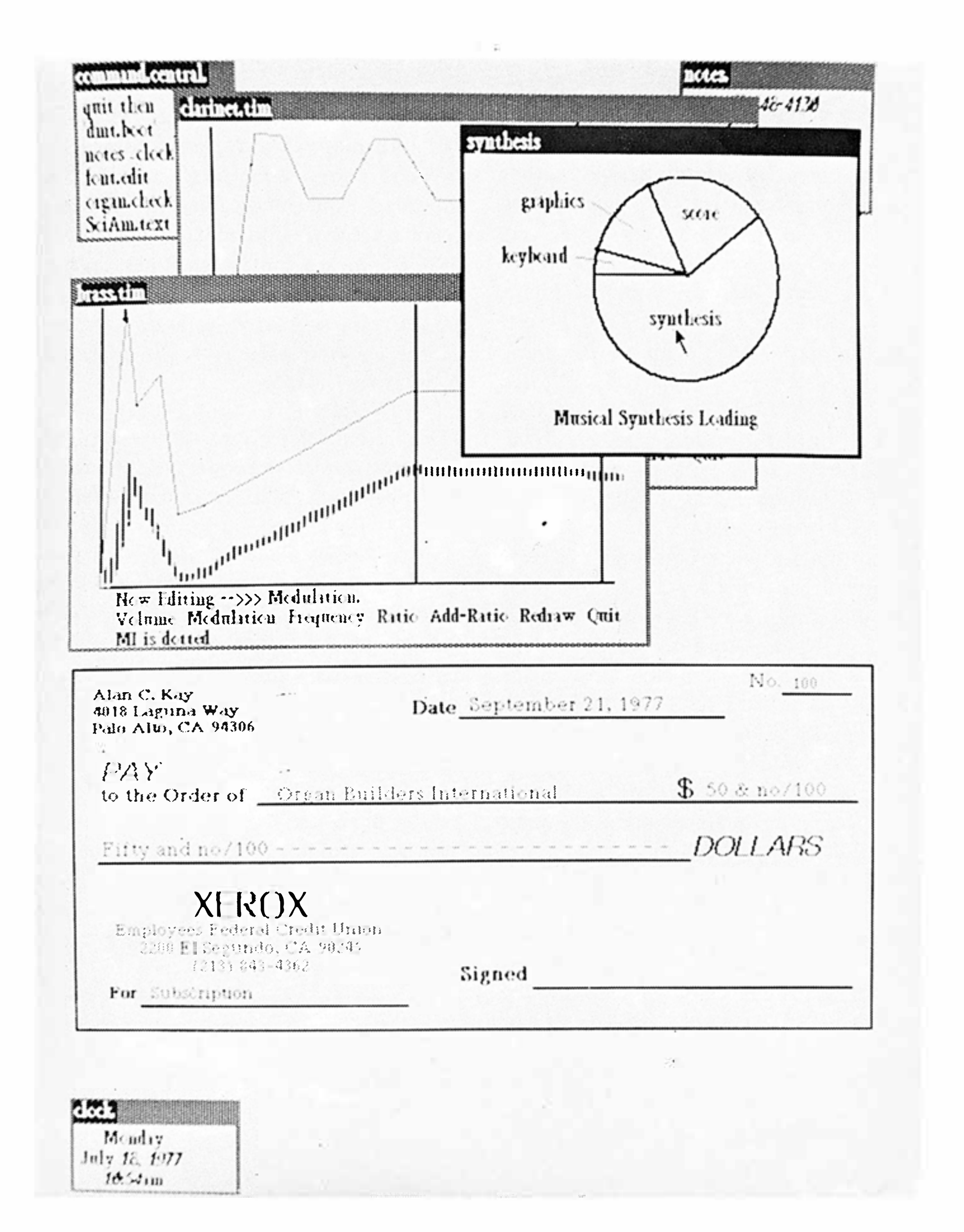
ORDENADOR EXPERIMENTAL, de uso personal, construido en el Xerox Palo Alto Research Center. Tiene por misión desarrollar un lenguaje de programación de alto nivel, que pudiera capacitar a los no expertos en la escritura de programas complicados. El autor y sus colegas se interesaron también en utilizar dicho ordenador para estudiar los efectos de los cálculos personales en el aprendizaje. La máquina es autosuficiente; consta de un teclado, un dispositivo indicador, un monitor de alta resolución y un sistema sonoro, todo ello conectado a una unidad de proceso y a una memoria de discos insertables. El monitor puede representar miles de caracteres de calidad semejante a los del material impreso.

nos dimos cuenta de que muchos de los problemas relacionados con el diseño de un ordenador de uso personal, particularmente aquellos que tenían algo que ver con la comunicación expresiva, se hicieron más patentes cuando se consideraron como usuarios a los niños menores, de 6 años. También observamos que los niños requerían más potencia por parte del ordenador de la que un adulto necesitaba en un sistema de tiempo compartido. Las mejores salidas que el tiempo compartido podía proporcionar eran líneas toscas dibujadas en un tono verde y tonos musicales de onda cuadrada. Pero los niños estaban acostumbrados a pintar con los dedos, a la televisión en color y a los discos estereofónicos, por lo que las cosas que podían hacer con un sistema en tiempo compartido de baja capacidad no les incitaban apenas ni lograban mantener su interés.

Debido a que el sistema LOGO no se diseñó para todas las personas y usos que habíamos pensado, decidimos no editarlo, sino idear otro nuevo sistema de programación que combinase la simplicidad y facilidad de acceso con una mejora cualitativa en la programación adulta a nivel de expertos. En ese esfuerzo nos guiamos, igual que en el sistema FLEX, por las ideas maestras del lenguaje de programación SIMULA, que desarrollaron a mediados de los años 60, Ole-Johan Dahl y Kristen Nygaard en el centro de cálculo noruego de Oslo.

Nuestro ordenador experimental es autosuficiente y encaja perfectamente en una mesa de trabajo. El almacenamiento de grandes cantidades de información corre a cargo de memorias de disco insertables que pueden contener el equivalente a 1500 páginas impresas de información (aproximadamente tres millones de caracteres). Aunque los monitores visuales de los años 80 serán, probablemente, mosaicos de pantalla lisa que reflejarán la luz de la misma forma que los relojes de cristal líquido, la mejor salida visual que hoy puede obtenerse es a través de un tubo de televisión de alta resolución, en blanco y negro o bien en color.

La salida sonora de alta fidelidad se produce por medio de una conversión interna de señales digitales discretas en formas de onda continuas, que luego se mandan a un audioamplificador y a los altavoces. El usuario realiza su entrada principal a través de un teclado parecido al de una máquina de escribir y un dispositivo indicador, llamado "cursor", que controla la posición de una flecha sobre la pantalla cuando se le pulsa desde la mesa que está junto al monitor. Entre otros sistemas de entrada, se encuentran un teclado parecido al de un órgano de música, un indicador en forma de lápiz, una palanca, un micrófono y una cámara de televisión.



LAS "VENTANAS", marcos del monitor dentro de la pantalla del monitor mayor, capacitan al usuario para organizar y editar información a diferentes niveles de claridad. Una vez creadas las ventanas, se solapan sobre la pantalla como hojas de papel. Cuando el dispositivo indicador selecciona una ventana parcialmente cubierta, ésta vuelve a visualizarse solapando a las otras ventanas. Se pueden visualizar simultáneamente imágenes con diversos grados de contenido simbólico. Estas comprenden: dibujos detallados, imágenes analógicas (gráficas) y simbólicas (números o palabras).

La actividad habitual de nuestro ordenador de uso personal consiste en la manipulación de simulaciones suministradas por el sistema SMALLTALK o creadas por el usuario. El estado dinámico de una simulación aparece en el monitor y su desarrollo general se modifica cuando el usuario cambia las imágenes pulsando órdenes en el teclado o apuntando con el "cursor". Por ejemplo, se simulan documentos de textos con diferentes formatos y tipos de impresión, de suerte que, en la pantalla, aparezca una imagen del documento. Este se edita apuntando a los caracteres y párrafos con el cursor, y luego borrando, añadiendo y reestructurando las partes del documento. Cada cambio se refleja instantáneamente en la imagen del documento.

En muchos casos, la pantalla del monitor resulta demasiado pequeña para contener toda la información que un usuario desee consultar de una vez; por eso es por lo que hemos desarrollado "ventanas" o marcos simulados del monitor, dentro del monitor físico mayor. Las ventanas organizan simulaciones para editar y mostrar información, permitiendo crear y visualizar, a diferentes niveles de claridad, un documento com-

puesto de textos, grabados, notas musicales, animaciones dinámicas, etcétera. Creadas las ventanas, éstas se solapan en la pantalla como hojas de papel; cuando el cursor apunta a una ventana parcialmente cubierta, ésta vuelve a visualizarse solapando a las otras ventanas. Estas ventanas que contienen información útil, pero no inmediatamente necesaria, se reducen a pequeños rectángulos que se etiquetan con un nombre, indicativo de la información que contienen. Un "toque" del cursor hace que se abran al instante y muestren sus contenidos.

En el estado actual de las cosas, el desarrollo del soporte lógico resulta mucho más difícil y consume más tiempo que el desarrollo del soporte físico. Con el tiempo, el ordenador de uso personal se compondrá de componentes electrónicos más o menos normalizados, pero el soporte lógico, que dará vida a las ideas del usuario, habrá de pasar por un largo y arduo proceso de refinamiento, si ha de servir, y no obstaculizar, a los fines a que está destinado un medio dinámico de uso personal.

Por esta razón hemos invitado a lo largo de los últimos cuatro años a unos 250 niños (de edades comprendidas entre los 6 y 15 años) y a 50 adultos, para probar las versiones del SMALLTALK y tantear formas de mejorarlo. Sus creaciones tan imaginativas y variadas como ellos mismos comprenden programas de contabilidad doméstica, almacenamiento y recuperación de información, enseñanza, dibujo, pintura, síntesis musical, escritura y juegos. Los diseños posteriores del SMALLTALK se vieron influidos y mejorados, en gran parte, por los proyectos de nuestros visitantes.

Cuando un niño o un adulto se encara por primera vez con un ordenador de uso personal, suele estar habituado en las actividades de su propia elección. Su primer impulso es utilizar el sistema para hacer cosas que ya ha hecho: el contable automatizará el papeleo y la contabilidad, el profesor mostrará las facetas gráficas y dinámicas del curso y el niño trabajará para crear dibujos y juegos. De hecho, la gente empieza, de forma natural, por concebir y construir herramientas personales. Aunque el hombre se ha caracterizado, a lo largo de la historia, como la especie creadora de instrumentos, ser los artesanos del utillaje ha sido incumbencia de los maestros técnicos. Una de las razones que explican ese hecho reside en que las tecnologías suelen exigir técnicas, materiales, herramientas y condiciones físicas especiales. Sin embargo, una peculiaridad importante de los ordenadores es la de que, para usarlos, cualquiera puede construir herramientas muy generales. Estas herramientas se fabrican a partir de los mismos materiales y con el mismo esfuerzo que otras creaciones más específicas.

En un comienzo, los niños se relacionan con nuestro ordenador pintando cuadros y dibujando líneas rectas sobre la pantalla del monitor con un apuntador en forma de lápiz. Luego, descubren que los programas pueden crear estructuras más complejas que las que ellos pueden hacer manualmente. Aprenden que un dibujo tiene varias representaciones, de las cuales la más obvia –que es la imagen– aparece sobre la pantalla. La representación más importante es el modelo simbólico y editable de un dibujo almacenado en la memoria del ordenador.

Por ejemplo, en un aparato ordenador, se puede construir la imagen de un camión a partir de modelos de ruedas, una cabina y una cama, siendo cada uno de estos elementos de un color distinto. A medida que los elementos simbólicos de un modelo se vayan modificando, su imagen irá cambiando igualmente en la pantalla.

Los adultos también aprenden a aprovechar las características del ordenador. Cierto artista profesional que nos visitó pasó varios meses construyendo diversas herramientas que se parecían a aquellas que había utilizado para esbozar las imágenes sobre el cartón. Llegó a descubrir que la pantalla de mosaicos (el almacenamiento indeleble, e instantáneamente borrable, del medio) y su nueva posibilidad de programar podían combinarse para crear ricas composiciones que no podían hacerse ni con lápiz ni con pincel. Partiendo del uso del ordenador para la simulación restringida de un medio ya existente, consiguió descubrir las posibilidades únicas del ordenador en el campo de la expresión humana.

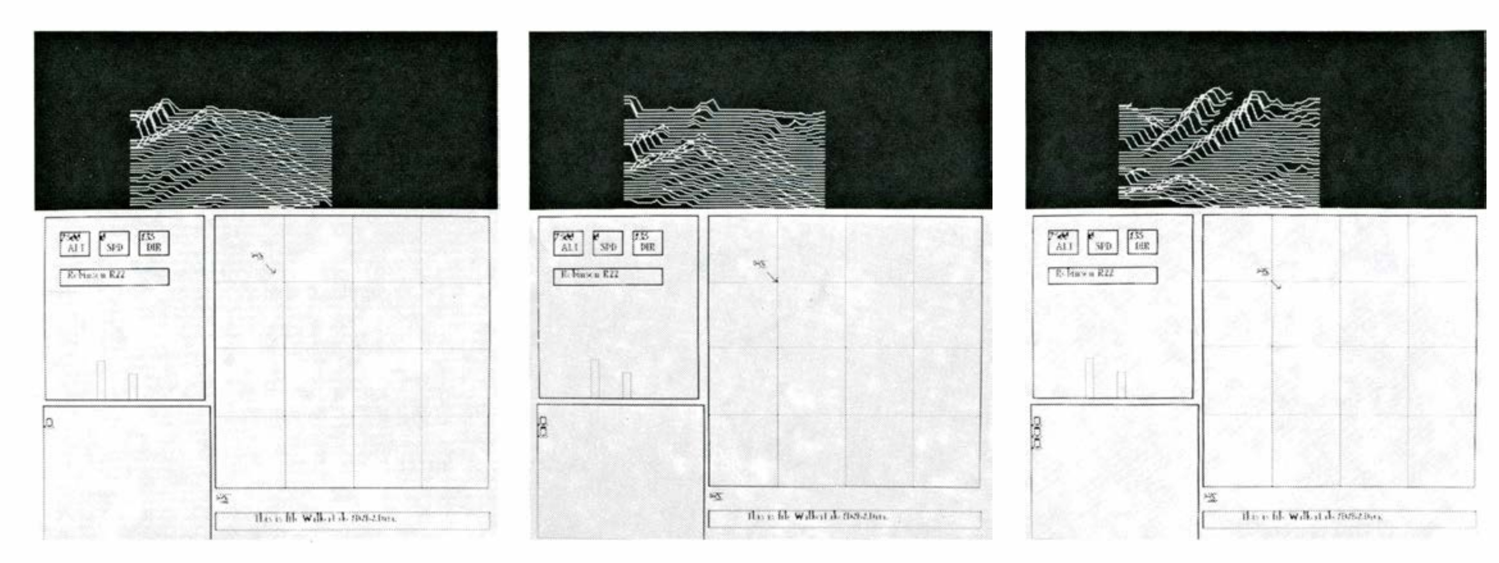
Una de las mejores formas de enseñar a los no expertos a comunicarse con ordenadores es hacerles ver los niveles de abstracción en los cuales se manejan las imágenes. La manipulación de imágenes sigue las etapas generales del crecimiento intelectual. Para un niño pequeño, una imagen es algo que crear, a saber, una mezcla libre de formas y colores desconectada del mundo real. Cuando son un poco mayores dibujan imágenes que representan directamente conceptos, verbigracia, personas, animales domésticos y casas.

trait name	description
name	box; picture; activity
location	□
angle	□
size	□
new	location ← center. angle ← 0. size ← 100.
show	✐ paint black. shape
erase	✐ paint background. shape
shape	✐ up. goto location; turn angle; down. 1 to 4 do (✐ go size; turn 90).
grow □	erase. size ← size + □. show

EL SMALLTALK es un nuevo lenguaje de programación desarrollado en el Xerox Palo Alto Research Center para emplearlo en el ordenador experimental de uso personal. Está compuesto de "actividades", entidades semejantes a ordenadores, que pueden realizar un conjunto específico de tareas y comunicarse con otras actividades dentro del sistema. Las nuevas actividades se crean enriqueciendo las familias de actividades existentes con "rasgos" adicionales, o propiedades que se definen con relación al método que ha de seguirse. La descripción de la familia "caja", que se muestra aquí, es un diccionario de sus rasgos. Para crear un nuevo miembro de la familia caja, se envía un mensaje al rasgo "nuevo" estableciendo las características de la caja nueva en términos de los valores específicos para los rasgos generales "posición", "ángulo" y "tamaño". En este ejemplo, se ha rellenado "nuevo" para especificar una caja localizada en el centro de la pantalla con un ángulo de cero grados y un lado de 100 puntos de largo. Para "mostrar" la nueva caja se dan direcciones a un miembro de la familia de dibujos curvos, "brocha", mediante el rasgo abierto "forma". Primero, la brocha se desplaza a la posición especificada. gira en la dirección apropiada y aparece en la pantalla. Luego, dibuja un cuadrado desplazándose la distancia indicada por "tamaño", girando 90 grados y repitiendo estas acciones tres veces más. El último rasgo de la lista está abierto, lo cual indica que un valor numérico va a ser proporcionado por el usuario cuando se requiera mediante un mensaje. Una caja "crece" borrándola primero, incrementando (o disminuyendo) su tamaño en el valor proporcionado por el mensaje y volviéndola a visualizar.

Interacción del mensaje	Efecto gráfico	Comentario
nueva caja llamada "joe"! caja: joe		Se crea un descendiente de la familia "caja" y se le llama "joe"
joe gira 30! ok		La caja joe recibe el mensaje y gira 30 grados
joe crece –15! ok		Joe se hace más pequeño en 15 unidades
joe bórrate! ok		Joe desaparece de la pantalla
joe muéstrate! ok		Joe reaparece
nueva caja llamada "jill"! caja: jill		Aparece una nueva caja
jill gira –10! ok		Sólo gira jill. Joe y jill son actividades independientes
1 a 10! intervalo: 1 2 3 4 5 6 7 8 9 10		Un intervalo representa una secuencia de números
siempre! intervalo: 1 2 3 4 5 6 7 8 9 10 11...		Siempre es el intervalo infinito. Debe terminarse golpeando una tecla de escape
haz de 1 a 10 (joe gira 20)! ok		Joe gira sobre sí mismo
haz siempre (joe gira 11, jill gira –13)! ok		Se crea una sencilla película paralela de joe y jill girando en direcciones opuestas mediante la combinación de siempre con una petición de giro tanto para joe como para jill.

SECUENCIA DE APRENDIZAJE SMALLTALK. Enseña a los estudiantes los consejos básicos del lenguaje haciéndoles interaccionar con una familia de actividades ya definida. Primero, se crean, se nombran y manipulan los descendientes de la familia caja, y se introduce una segunda familia llamada "intervalo". Luego, los descendientes de "intervalo" y caja se combinan para generar una animación de dos cajas girando.



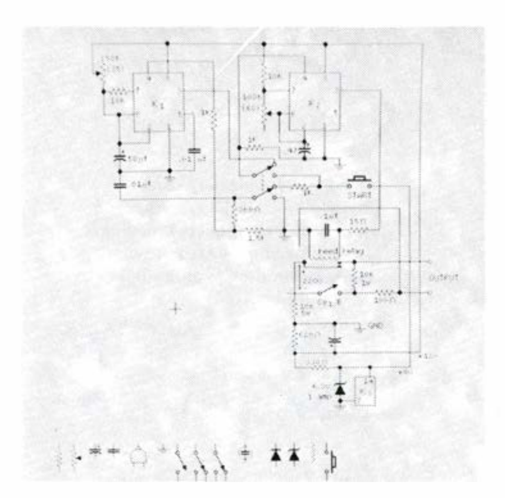
SIMULACION DE UN HELICOPTERO, desarrollada por un estudiante de 15 años. El usuario dirige la ruta del helicóptero con el dispositivo indicador, que controla la posición de la flecha negra sobre la pantalla. La ventana superior muestra la topografía cambiante a me-

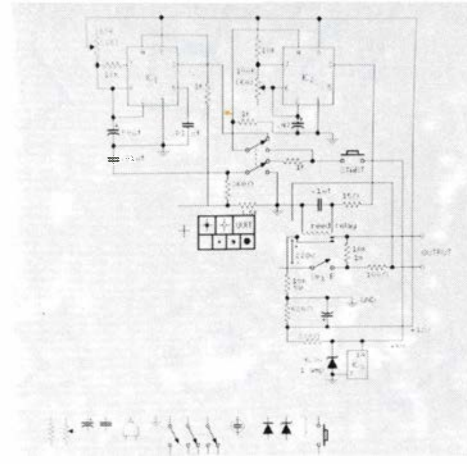
Más tarde, aparecerán las imágenes analógicas cuyas formas están estrechamente relacionadas con su significado y utilidad, tales como las figuras geométricas y gráficas. Finalmente, las imágenes simbólicas se utilizan para representar conceptos que son demasiado abstractos para razonar analógicamente, por ejemplo, números, términos lógicos y algebraicos, y los caracteres y palabras que constituyen el lenguaje.

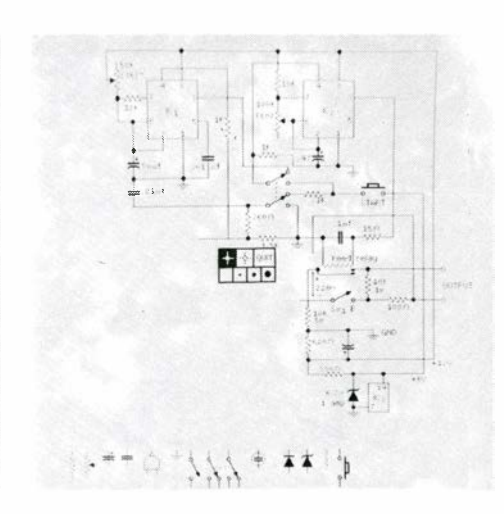
Los tipos de imagen de esta jerarquía son cada vez más difíciles de representar en un ordenador. Pueden dibujarse o pintarse fácilmente formas libres e imágenes literales mediante líneas y semitonos en la matriz de puntos de la pantalla del monitor, con la ayuda del cursor, o en unión de programas que dibujen curvas, rellenen superficies con diferentes tonalidades o muestren perspectivas de modelos tridimensionales. También se pueden generar imágenes analógicas, por ejemplo, el modelo de un instrumento musical simulado: un gráfico secuencial que represente la evolución dinámica de la amplitud, la variación de frecuencias y la gama tonal.

La utilidad de las representaciones simbólicas se manifiesta en que ofrecen un medio para manejar conceptos que son difíciles de figurar directamente, como las generalizaciones y las relaciones abstractas. Además, a medida que una imagen gana en complejidad, su propiedad más importante, la de hacer instantáneamente claras las relaciones locales, pierde en utilidad. La comunicación con ordenadores basada en símbolos, como suele ocurrir en el lenguaje natural, ha resultado ser mucho más difícil de lo que muchos habían pensado. La razón está en no entender cómo los seres humanos recurren al contexto de su experiencia para interpretar las ambigüedades de un discurso común. Si todavía no se entiende cómo los seres humanos hacen lo que hacen, habrá que esperar tiempo para comprometer a los ordenadores en acciones parecidas. Ello no obsta que se puedan idear lenguajes artificiales para un ordenador, los cuales representen conceptos y actividades que nosotros entendemos, y que tengan una estructura básica lo suficientemente sencilla como para que puedan ser fácilmente aprendidos y utilizados por los no expertos.

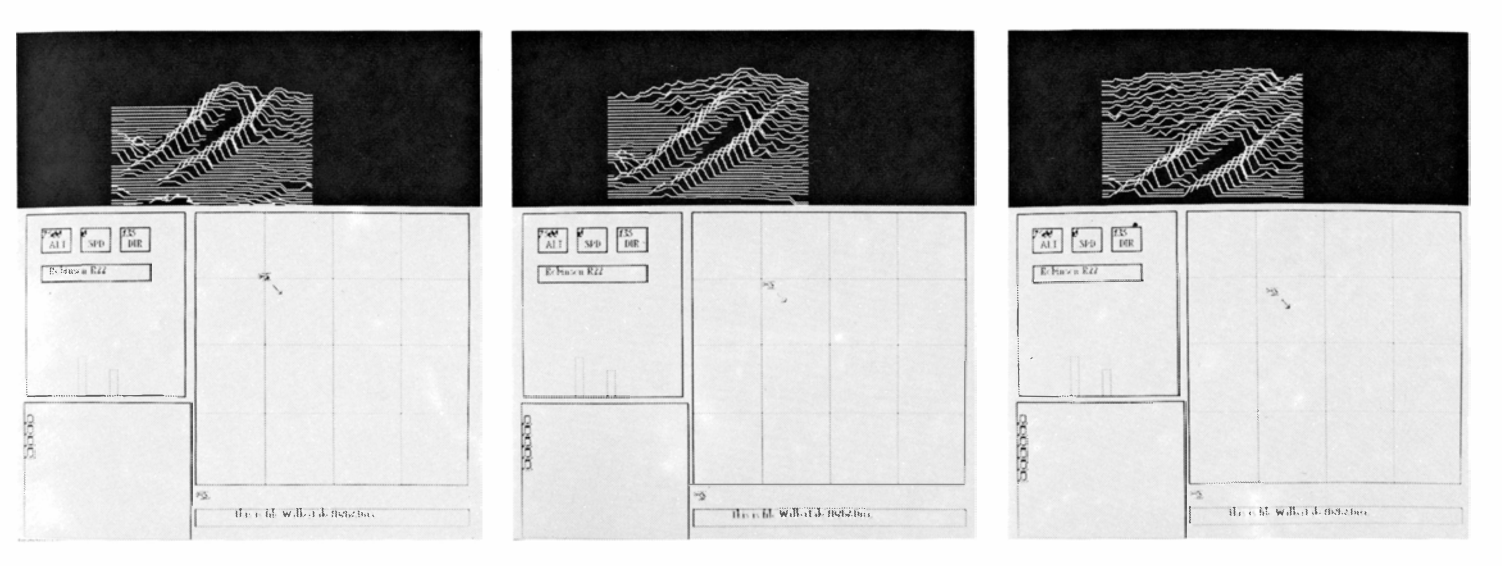
La estructura particular de un lenguaje simbólico es importante porque aporta un contexto en el cual algunas ideas se pueden imaginar y expresar con mayor soltura que otras. Por ejemplo, la notación matemática surgió por primera vez para abreviar conceptos que sólo podían







PROGRAMA PARA DIBUJAR CIRCUITOS, desarrollado por un muchacho de 15 años. Permite al usuario construir un diagrama complejo de circuitos seleccionando las componentes a partir de una "lista" visualizada en la parte inferior de la pantalla. Luego, los componentes



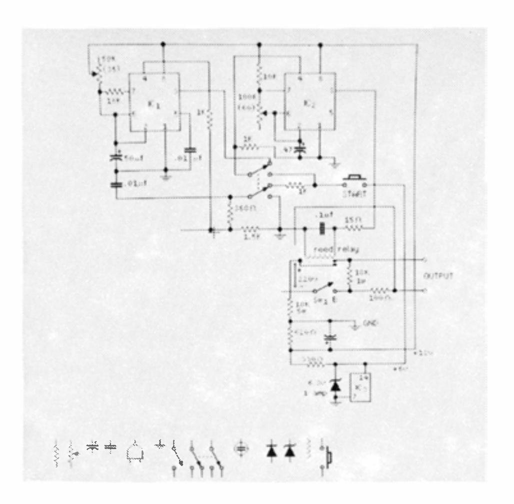
dida que vuela el helicóptero. (Los terrenos reales se obtuvieron de los mapas Landsat.) Una tercera ventana detecta la altitud, dirección y velocidad del helicóptero. La variedad de sucesos que se pueden simular simultáneamente demuestra la potencia del procesamiento paralelo.

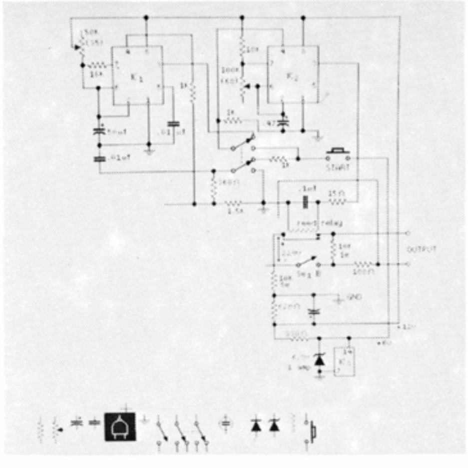
expresarse con torpes circumloquios en un lenguaje natural. Poco a poco se fue comprendiendo que la forma de una expresión podía ser de gran utilidad en la concepción y manipulación del significado que dicha expresión representaba. Se dio un gran paso al idearse una nueva notación para representar conceptos que no encajaban en absoluto en la herencia lingüística cultural, verbigracia: planos funcionales, índices continuos y límites.

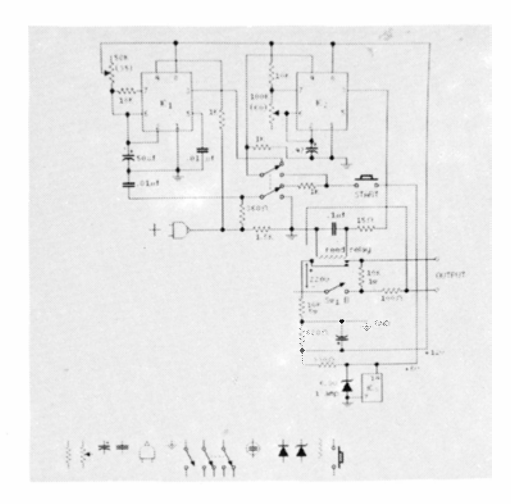
El ordenador creó nuevas necesidades de lenguajes invirtiendo el proceso tradicional de la investigación científica. Esto trajo como consecuencia la configuración teórica de nuevos universos para producir fenómenos simulados. De acuerdo con ello, se necesitaron estructuras simbólicas para comunicar conceptos tales como descripciones imperativas y estructuras de control.

La mayoría de los lenguajes de programación que se encuentran hoy en servicio se desarrollaron como medios simbólicos para manejar los conceptos a nivel de soporte físico de los años 50. Este enfoque nos llevó a dos tipos de bloques pasivos: estructuras de datos, o materiales de construcción inertes, y procedimientos, o algoritmos paso a paso para manipular datos. Los lenguajes basados en estos conceptos (entre ellos, BASIC, FORTRAN, ALGOL y APL) siguen sus descripciones de una forma estrictamente secuencial. Puesto que un dato puede sufrir un cambio por cualquier procedimiento que lo encuentre, el programador ha de tener sumo cuidado en seleccionar únicamente aquellos procedimientos que sean los apropiados. A medida que se acometan sistemas más complejos, sistemas que requieran complicadas combinaciones de procedimientos, la dificultad en conseguir que todo el sistema funcione aumentará geométricamente. Aunque a la mayoría de los programadores se les sigue enseñando los lenguajes de procedimiento de datos, existe ya un común acuerdo sobre su insuficiencia.

Hay otro enfoque, más prometedor; consiste en idear bloques de mayor generalización. Los datos y procedimientos se reemplazan por la idea única de "actividades", esto es: entidades semejantes a ordenadores que se manifiestan cuando reciben un mensaje apropiado. Dicho lenguaje carece de sustantivos y verbos, sólo existen actividades dinámicamente comunicables. Cada transacción, descripción y proceso de control se entienden en términos de mensajes que se envían y se reciben de las actividades del sistema. Es más, cada actividad pertenece a una familia de actividades semejantes; y este grupo posee la capacidad de reconocer y responder a mensajes dirigidos







se ubican y conectan mediante el dispositivo indicador. Se puede generar una lista adicional pulsando un botón en el dispositivo indicador; esta lista proporciona puntos abiertos y vacíos y líneas de diferente anchura. En la secuencia se seleccionan y se añaden dos componentes.



ANIMACION DE UNA CARRERA DE CABALLOS, que muestra las posibilidades del ordenador de uso personal para crear imágenes dinámicas. El rango posible de tales simulaciones está limitado por la versatilidad del lenguaje de programación y la imaginación del usuario. Las imágenes de caballos, jinetes y fondo se recuperan de forma independiente, desde los archivos de almacenamiento, y se ordenan con el dispositivo indicador para la simulación de la carrera. Una sola orden tecleada origina que los dos caballos y jinetes compitan a través de la pantalla.

a las actividades, así como ejecutar determinadas acciones específicas, a saber: dibujar ilustraciones, producir sonidos o sumar números. Mediante la combinación y el enriquecimiento de los "rasgos" o las propiedades heredadas de las familias existentes, se crean otros nuevos grupos o familias.

Un sistema "mensaje-actividad" es esencialmente paralelo: cada actividad está constantemente preparada para enviar y recibir mensajes. Ocurre pues que el ordenador patrón se halla realmente dividido en miles de ordenadores, cada uno con las capacidades del conjunto. Por tanto, el enfoque mensaje-actividad permite la representación dinámica de un sistema a muchos niveles de organización, desde el atómico al macroscópico, pero con una "capa" de protección en cada nivel cualitativo de detalle a través de la cual deben enviarse y comprobarse los mensajes de trabajo. Este nivel de complejidad puede manejarse con seguridad ya que el lenguaje limita rigurosamente los tipos de interacciones entre actividades, permitiendo sólo las apropiadas, lo mismo que una hormona únicamente puede intervenir en las células de respuesta específica. El SMALLTALK, que es el sistema de programación de nuestro ordenador de uso personal, constituyó el primer lenguaje basado íntegramente en los conceptos estructurales de mensajes y actividades.

El tercer, y más reciente, marco para la comunicación de alto nivel es el lenguaje "observador". Aunque los lenguajes mensaje-actividad representan un avance respecto a la estructura de procedimiento de datos, las relaciones entre las diversas actividades son de algún modo independientes y analíticas. No obstante, muchos conceptos se hallan entrelazados con tal densidad que el análisis les obliga virtualmente a desaparecer. Por ejemplo, es sabido que la física del siglo XX concede igual importancia al fenómeno y a su contexto, ya que distintos observadores con diferentes puntos de referencia perciben el mundo de modo diverso. En un lenguaje observador, las actividades se sustituyen por "puntos de vista" que se entrelazan para formar correspondencias entre conceptos. Por ejemplo, un perro puede ser visto abstractamente (como un animal), analíticamente (como un ser compuesto de órganos, células y moléculas), pragmáticamente (como vehículo para un niño), alegóricamente (como un ser humano en un cuento de hadas) y contextualmente (como un medio de fertilizar un césped). Nos hallamos en el momento en que empiezan a formularse los "lenguajes observador". Ellos y sus sucesores serán los vehículos de comunicación de los años 80.

Nuestra propia experiencia, y la de otros que enseñan programación, es que el estilo particular y los conceptos principales de un primer lenguaje de ordenador, no sólo ejercen una fuerte influencia sobre lo que un nuevo programador pretende conseguir, sino que, además, dejan una huella, en la programación y en el ordenador, que puede durar años. El proceso de aprendizaje para programar un ordenador puede imponer un punto de vista tan particular, que llegue a resultar frustrante, para el nuevo programador, aprender otros métodos alternativos de percepción y resolución de problemas.

Al principio de nuestro estudio consideramos tímidamente las características de simulación de los lenguajes de procedimiento de datos que los niños habían sido capaces de aprender, como el BASIC y LOGO. Luego, preocupados porque el proceso de impresión pudiera perjudicar la absorción de ideas más fuertes, decidimos buscar un modo de presentar las ideas mensaje-actividad del SMALLTALK en términos concretos, sin ambigüedad. Esto lo hicimos empezando con situaciones sencillas que incorporaban un concepto; luego, fuimos incrementando gradualmente la complejidad de los ejemplos para llevar el concepto a su plena generalidad. Aunque el modelo tipo comunicación del SMALLTALK es un modo bastante abstracto de representar descripciones, vimos con asombro que los grupos de niños que lo utilizaban parecían encontrar las ideas tan fáciles de aprender como las de lenguajes más concretos.

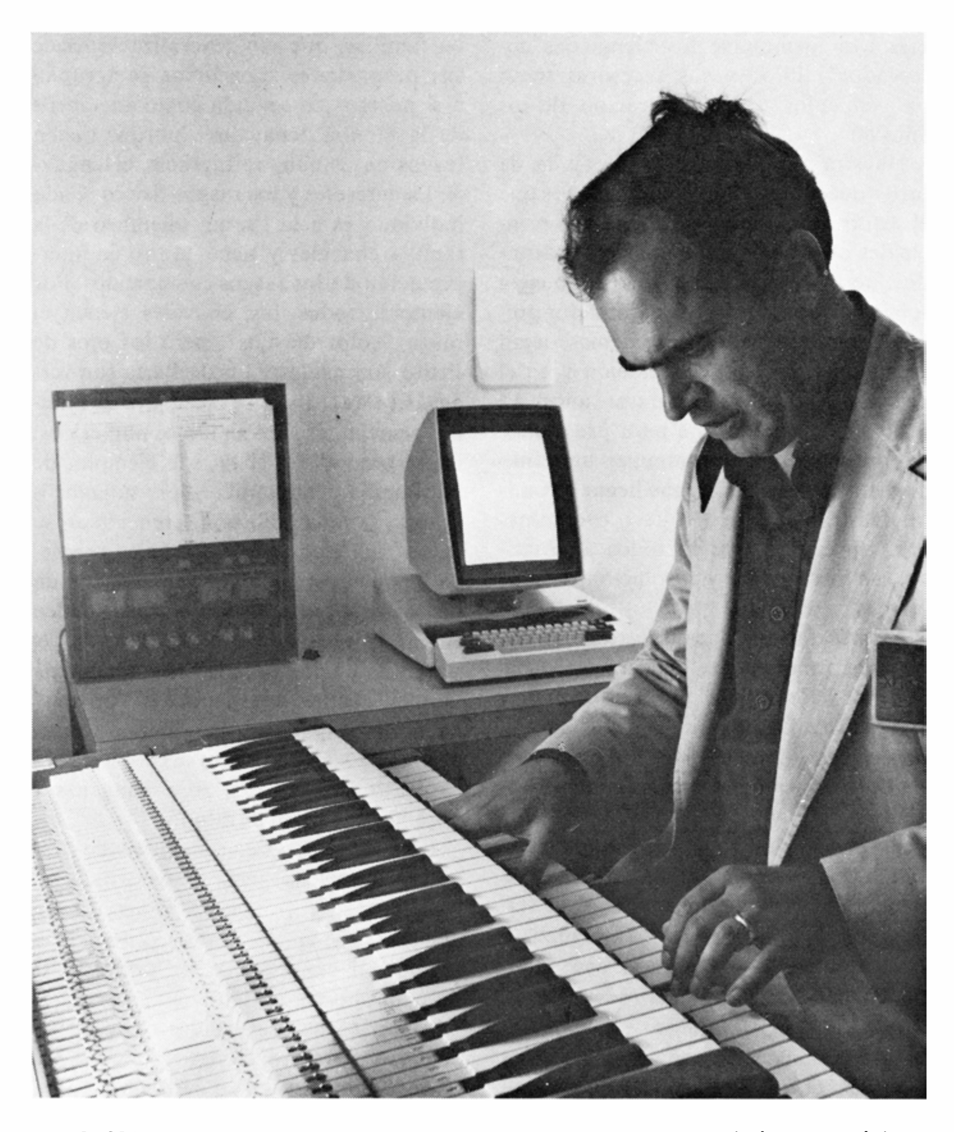
Por ejemplo, debido a que la mayoría de los lenguajes de programación pueden ocuparse de una sola cosa a la vez es difícil representar con ellos situaciones tan simples como niños en un colegio, naves espaciales en el cielo o balones botando en el suelo. En el SMALLTALK se manejan desde un principio los modelos paralelos, y los niños no parecen tener mayor dificultad en utilizarlos. El proceso paralelo resulta estrechamente afín a la vía que sigue el pensamiento. Cuando se camina por una calle, parte del cerebro puede pensar en la ruta, parte en la hora de cenar y otra tercera parte tal vez esté admirando la puesta de sol, y así sucesivamente.

Otra característica importante del SMALLTALK es la clasificación de objetos en familias, que son generalizaciones de sus propiedades. Los niños se agrupan a sí mismos en seguida como miembros de la familia "chavales", porque tienen rasgos en común, verbigracia, el lenguaje, los intereses y los rasgos físicos. Cada individuo es a la vez un miembro de la familia chavales y tiene su propia interpretación de los rasgos compartidos. Por ejemplo, todos los chavales tienen el rasgo "color de ojos", pero los ojos de Pedro son azules y los de Berta son verdes. El SMALLTALK se construye de familias semejantes. Los símbolos numéricos, tales como el 2 y el 17, son ejemplos de la familia "número". Los miembros de esta familia sólo se diferencian en su valor numérico (que es su única propiedad) y comparten una definición común de los diferentes mensajes que pueden enviar o recibir. En el SMALLTALK, el símbolo de una "brocha" es también una familia. Todos los símbolos "brocha" tienen la capacidad de dibujar líneas, pero cada símbolo posee su propio conocimiento de cuál sea su orientación y posición en el área de dibujo.

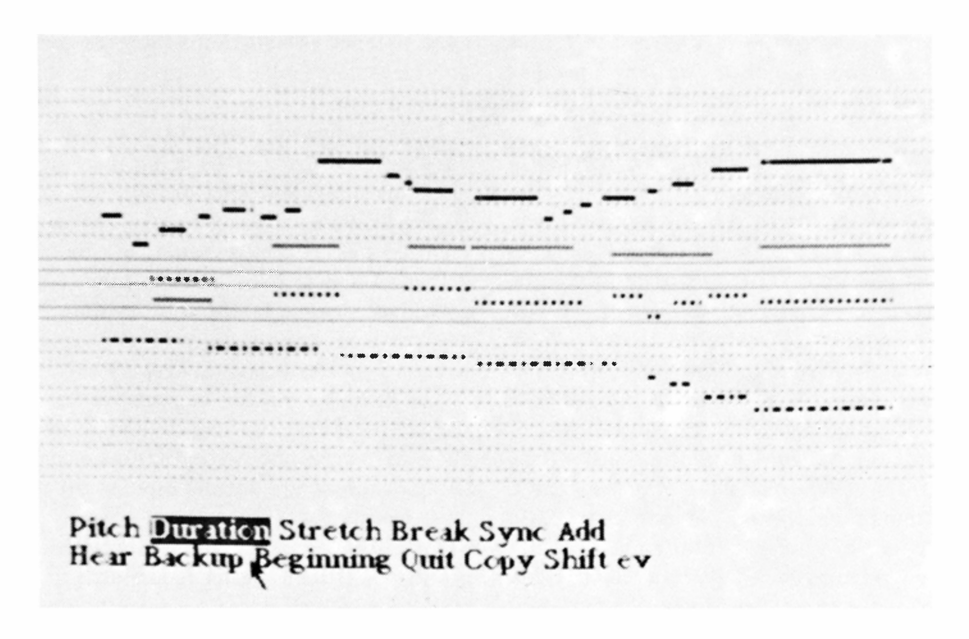
La descripción de un lenguaje de programación viene generalmente dada en función de su gramática: el significado que se supone que cada construcción gramatical expresa y el método utilizado para obtener el significado. Por ejemplo, varios lenguajes de programación emplean construcciones gramaticales tales como (PLUS 3 4) o 3 ENTER 4 + para especificar la intención de sumar el número 3 al número 4. Estas frases tienen el mismo significado. En el ordenador, cada construcción debe dar lugar al número 7, aunque los métodos empleados para obtener la solución se diferencien considerablemente de un tipo de ordenador a otro.

La gramática del SMALLTALK es sencilla y precisa. Cada frase es un mensaje para una actividad. Una descripción de la actividad deseada viene seguida de un mensaje que selecciona un rasgo de la actividad que se hace llevar a cabo. La actividad designada decidirá si quiere aceptar el mensaje (que es lo que suele hacer) y posteriormente actuará sobre él. Puede haber muchos mensajes concurrentes pendientes de una actividad, incluso para el mismo rasgo. El transmisor del mensaje puede decidir esperar o no una respuesta. Generalmente espera, pero puede decidir emprender otros asuntos si el mensaje ha recurrido a un método que precisa de un cálculo muy laborioso.

La integración de los conceptos del lenguaje de programación con los con-



La MUSICA puede representarse en un ordenador de uso personal como imágenes analógicas. Las notas tocadas en el teclado se "transforman" en un pentagrama secuencial en el monitor.



PENTAGRAMA MUSICAL, que se generó mientras se tocaba la música en el teclado. La notación simplificada representa el tono, por la situación vertical, y la duración, por la longitud horizontal. Las notas se pueden acortar, alargar o cambiar, y la pieza modificada se repite luego.

ceptos de edición, gráficas y recuperación de información, deja disponible una amplia gama de actividades útiles a las que el usuario puede recurrir con poco o ningún conocimiento de programación. Los estudiantes aprenden a manejar el SMALLTALK enviando mensajes a familias de actividades ya existentes, tales como la familia "caja", cuyos miembros se muestran en la pantalla como cuadrados. Una caja puede cambiar aisladamente su tamaño, posición, rotación y forma. Después de algunas experiencias con el envío de mensajes que originan efectos sobre la pantalla del monitor, el estudiante puede echar un vistazo a la definición de la familia caja. En el SMALLTALK cada familia se describe con un diccionario de rasgos que se definen con arreglo al método que ha de seguirse. Por ejemplo, el mensaje "joe crece 50" significa: encontrar la actividad de nombre "joe", encontrar su rasgo general llamado "crece __" y rellenar su parte vacía con el valor específico 50. El estudiante puede añadir fácilmente un rasgo nuevo, análogo a los ya presentes en la definición de la familia (tales como "crecer" o "girar"). La siguiente fase de aprendizaje incluye una elaboración de este tema básico mediante la creación de juegos: guerras espaciales y útiles para dibujar y pintar.

Hay dos enfoques fundamentales para el cálculo personal. El primero, semejante a la improvisación musical, es de carácter exploratorio: se producen los efectos para ver cómo son; se persiguen los errores, se entienden y se concretan. El segundo enfoque, que se parece a la composición musical, requiere bastante más planificación, generalidad y estructura. Se utiliza el mismo lenguaje para ambos métodos, aunque el marco es totalmente diferente.

De nuestro estudio hemos aprendido la importancia de un equilibrio entre la libre exploración y el curso planificado. La experiencia en cálculo personal se parece a la introducción de un piano en un aula de tercer grado. Los niños harán ruido e incluso llegarán a tocar algo, para probar; con el tiempo, sin embargo, necesitarán ayudar para utilizar el instrumento de un modo pertinente. También hemos descubierto que los diferentes niveles de abstracción proporcionados por el SMALLTALK no son igualmente accesibles para los niños. La idea central de la simbolización se basa en dar un nombre sencillo a un conjunto complejo de ideas y recurrir, posteriormente, a las mismas a través de dicho nombre. Al

observar a bastantes niños, de edades comprendidas entre los 6 y 7 años, que habían conseguido dar este paso en sus programas, vimos que sus capacidades para predecir y visualizar las consecuencias de las acciones que ellos podían tomar estaban limitadas.

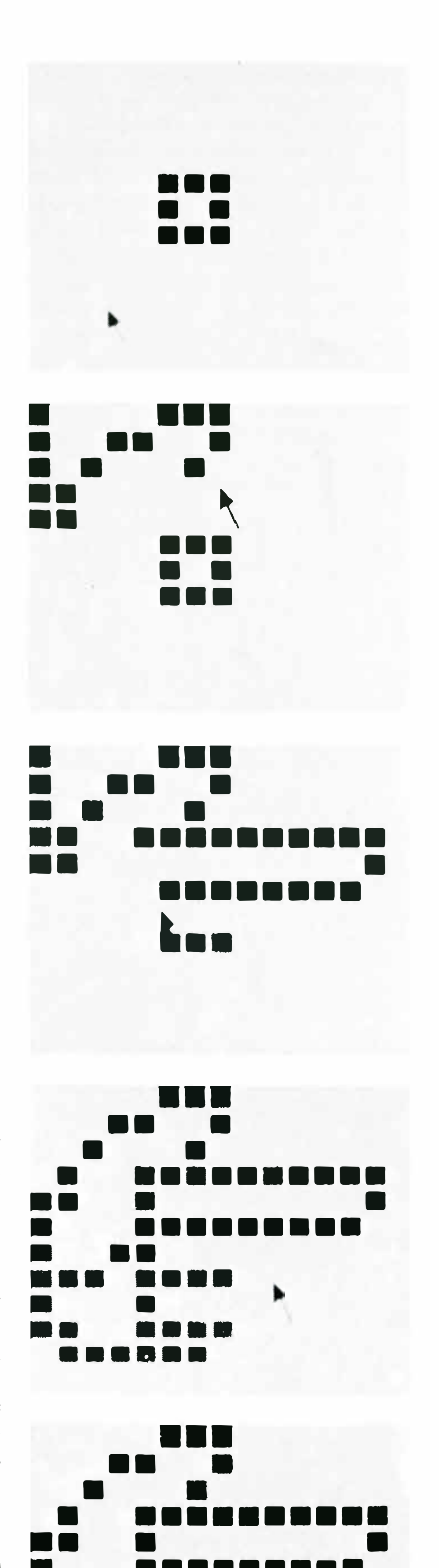
Los niños de edades comprendidas entre los 8 y 10 años poseen un desarrollo gradual de su capacidad para visualizar y planificar: pueden utilizar, asimismo, el concepto de familias y una forma más sutil de nomenclatura, esto es, el uso de ciertos rasgos, como el tamaño, que pueden representar diferentes valores numéricos en tiempos distintos. No obstante, para la mayoría de los niños, las verdaderas implicaciones de una generalidad simbólica más amplia no resultan obvias en absoluto. Cuando alcanzan los 11 y 12 años, registramos una considerable mejoría en su capacidad para idear estructuras generales e ingeniar herramientas con la totalidad del ordenador. Los adultos superan las etapas más rápidamente que los niños, y se suelen mostrar capaces de crear herramientas a las pocas semanas de práctica. No se sabe si las etapas del desarrollo intelectual observadas en los niños están relativa o absolutamente relacionadas con la edad; cabe que, por estar expuestos a un mundo en el que la creación simbólica se ve recompensada por efectos maravillosos, se pueda reducir el tiempo requerido por los niños para madurar de una etapa a otra.

La limitación más importante para los no expertos en el cálculo personal aparece cuando conciben un proyecto que, aunque es fácil de hacer en el lenguaje, requiere conceptos de diseño que aún no han asimilado. Por ejemplo, es fácil construir un puente de ladrillos si se domina la idea de arco, pero de otra forma sería difícil o imposible. Evidentemente, a medida que aumenta la complejidad, la "arquitectura" domina al "material". La necesidad de métodos para caracterizar y comunicar conceptos arquitectónicos en el desarrollo de programas ha constituido un problema desde hace mucho tiempo en el diseño de sistemas de cálculo. Un lenguaje de programación proporciona un contexto para desarrollar estrategias y debe aportar un estilo que sugiera enfoques útiles para dar vida a los conceptos, amén de la posibilidad de crear instrumentos.

Tenemos la seguridad, basada en nuestra experiencia, de que los ordenadores de uso personal llegarán a ser una parte esencial de la vida humana de los años 80. A casi todo el mundo habrá de importarle la edición, el ahorro y la selección de cualquier tipo de información. Podemos comparar a las formas más complejas de cálculo a la música, en el sentido de que la mayoría de la gente llegará a conocerlas y gozarlas, y sólo unos cuantos alcanzarán a verse implicados en ellas directamente.

¿Cómo afectarán los ordenadores de uso personal a la sociedad? La interacción de la sociedad y el nuevo medio de comunicación y autoexpresión puede ser inquietante, incluso cuando la mayoría de los miembros de la sociedad aprendan a utilizar el medio rutinariamente. Los efectos sociales e individuales del nuevo medio son sutiles, y nada fáciles de percibir por la sociedad y el individuo. Vale la pena remontarnos a lo ocurrido con la escritura y extrapolarlo en las posibles reacciones ante la presentación del nuevo medio: analfabetismo, alfabetización y creación artística. Una vez disponible el material de lectura, fueron analfabetos aquellos que se quedaron atrás con respecto al nuevo medio. Resultó inevitable que unos cuantos individuos, creadores, emplearan la palabra escrita para expresar sus ideas y pensamientos íntimos. Los cambios más profundos se debieron a los cultos. No eran, forzosamente, mejores personas o mejores miembros de la sociedad, pero llegaron a considerar el mundo de una forma totalmente diferente de cómo lo habían contemplado antes, con consecuencias que fueron difíciles de predecir o controlar.

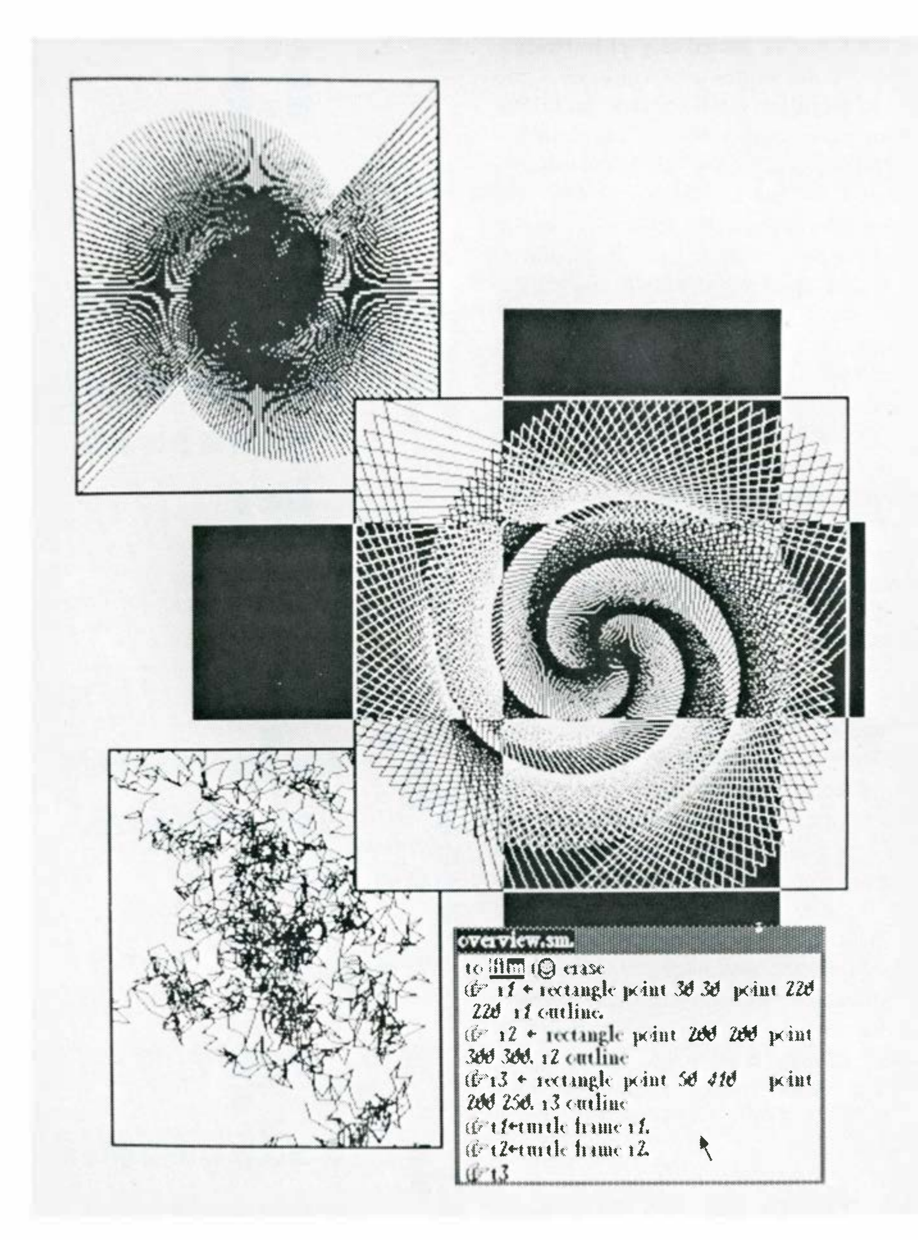
Es de esperar que los cambios originados en la cultura por el ordenador lleguen tan lejos como los que se dieron con la lectura y la escritura. Ahora bien, para la mayoría de las personas los cambios serán sutiles y no necesariamente en la dirección de sus esperanzas idealizadas. Por ejemplo, no deberíamos predecir o esperar que el ordenador de uso personal alumbrara una nueva revolución en la educación, sólo porque pueda hacerlo. En este siglo, cada nuevo medio de comunicación (el teléfono, el cine, la radio y la televisión) ha suscitado parecidas



LAS FUENTES DEL MONITOR pueden diseñarse sobre el ordenador de uso personal, a partir de una matriz de cuadrados blancos y negros. Cuando las fuentes se reducen, se semejan, en cuanto a calidad se refiere, a las del material impreso. Esta imagen de una mano punteada es un símbolo en SMALLTALK que representa el concepto de una palabra literal, verbigracia, el nombre asociado a una actividad.

predicciones que nunca se cumplieron. Se cuentan por millones las personas incultas del mundo que tienen fácil acceso a la cultura acumulada de siglos, en las bibliotecas públicas, y no hacen uso de ella. No obstante, una vez que el individuo o la sociedad decidan que la educación es esencial, el ordenador de uso personal, como antes el libro, figurará entre los vehícuos principales de que dispone la comunidad para la transmisión del saber.

También debe considerarse el impacto social de la simulación (la propiedad central del cálculo). Primero, como ocurre con el lenguaje, el usuario del ordenador se siente fuertemente instigado a acentuar la semejanza entre la simulación y la experiencia e ignorar las grandes distancias que los símbolos interponen entre los modelos y el mundo real. Son frecuentes los sentimientos de poder y fascinación narcisista con la imagen de uno mismo reflejada por la máquina. Otras tendencias adicionales son la de emplear el ordenador en misiones superficiales (simulando lo que el papel, la pintura y el archivo pueden hacer), utilizarlo como una "muletilla" (es decir, usarlo para registrar cosas que podemos perfectamente recordar nosotros mismos) o como una excusa (culpando al ordenador los fallos humanos). Más grave es la propensión humana de fiar y asignar altos poderes a un objeto que no se comprende totalmente. El hecho de que realmente muchas organizaciones basen sus decisiones en –y lo que es peor, tomen sus decisiones a partir de– ordenadores, es profundamente inquietante dado el estado actual de la informática. Sentimientos semejantes a los que produce la palabra escrita persisten todavía: si algo está "en blanco y negro" de alguna forma ha de ser verdad.



DIBUJOS COMPLEJOS, que pueden obtenerse en el ordenador de uso personal con descripciones muy concisas en SMALLTALK. Se realizan por medio de la repetición, rotación, reducción, superposición y combinación de dibujos de formas geométricas sencillas. En primer lugar, los estudiantes que aprenden a programar, crean imágenes interesantes de forma libre o imágenes literales, dibujándolas directamente en la matriz de puntos de la pantalla del monitor. Con el tiempo, aprenden a emplear las imágenes simbólicas en un lenguaje de programación para dirigir al ordenador hacia la generación de imágenes más completas que las que pudieran realizar a mano.

Los niños que no han perdido todavía gran parte de su capacidad de maravillarse y divertirse, nos han ayudado a encontrar un comportamiento en lo referente a nuestra relación con los ordenadores: no hay que automatizar el trabajo en el que se está involucrado, sólo los materiales. Si a uno le gusta dibujar, no debe automatizar el dibujo, sino programar el ordenador de uso personal para proporcionar un nuevo juego de pinturas. Si a uno le gusta la música, no hay que construir un "piano"; es mejor programar un nuevo tipo de instrumento.

Una falacia popular con respecto a los ordenadores es que son lógicos. Sinceros los define con mayor precisión. Como los ordenadores pueden contener descripciones arbitrarias, puede llevarse a cabo cualquier conjunto de reglas, sean o no coherentes. Además, el uso de símbolos en los ordenadores, como el uso de símbolos en el lenguaje y las matemáticas, está suficientemente desconectado del mundo real para permitir la creación de sublimes tonterías. Aunque el soporte físico del ordenador está sujeto a leyes naturales (los electrones sólo pueden moverse en los circuitos según ciertos caminos físicamente definidos), la gama de simulaciones que el ordenador puede realizar solamente está restringida por los límites de la imaginación humana. En un ordenador se puede conseguir que las naves espaciales viajen más deprisa que la velocidad de la luz, y que el tiempo retroceda.

Sólo en el caso de que creamos que lo que sabemos es acertado y completo, puede parecer absurdo hablar de la simulación de tonterías. Pero la historia no ha complacido a quienes compartieron este punto de vista. Por esto justamente es por lo que este mundo de tonterías aparentes ha de ser accesible a las futuras mentalidades en desarrollo. Aunque al ordenador de uso personal se le puede hacer trabajar como queramos, ¡sería un verdadero fallo hacerle actuar como una máquina!

Juegos matemáticos

Secciones cónicas, superficies regladas y otras manifestaciones de la hipérbola

Martin Gardner

Lewis Carroll escribió en cierta ocasión (en *The Dynamics of a Particle*): "¿Qué matemático que haya meditado largamente acerca de la hipérbola, que haya despedazado la desdichada curva con líneas de intersección acá y allá, en sus esfuerzos por demostrar alguna propiedad, tal vez mera calumnia, después de todo, no ha terminado por imaginar que el sufrido lugar geométrico está extendiendo sus asíntotas en silencioso reproche, o que le está guiñando pícaramente un foco con despectiva compasión?".

En una habitación oscura, pongamos una bola grande, por ejemplo, un balón de fútbol, sobre una mesa de color claro. Iluminemos la pelota directamente desde arriba con una linterna o lámpara de mano, como se muestra en la parte *a* de la ilustración de la página opuesta. La sombra de la pelota es, evidentemente, un círculo, cuyo centro es el punto de reposo del balón sobre la superficie de la mesa.

Movamos la linterna hacia la derecha, como se muestra en la parte *b* de la ilustración; la sombra se alarga, convirtiéndose en una elipse. El centro del círculo se ha descompuesto en dos puntos, que son los focos de la elipse. La esfera reposa sobre el foco más cercano a la fuente de luz. Conforme la luz se desplaza más y más hacia la derecha, el otro foco se mueve hacia la izquierda, aumentando la excentricidad de la elipse.

Bajemos la fuente de luz hasta enrasarla con la cima de la pelota *(c)*. El balón reposa todavía sobre el foco de la derecha, pero ahora el foco izquierdo ha viajado hasta el infinito. El perfil de la sombra es una parábola.

Hagamos descender la fuente de luz hasta que se encuentre más baja que la cima del balón *(d)*. La curva delimitada por la sombra se transforma en una hipérbola. La esfera todavía toca a la mesa en el foco de la hipérbola, pero ahora al foco desaparecido le ha ocurrido algo fantástica.

Imaginemos que las cuatro figuras tuvieran una "contrabola" idéntica a la esfera de la mesa y situada a igual distancia de la fuente luminosa, aunque del otro lado. Dicha contrabola se señala con línea de trazos en la última figura *(d)*. Observemos que proyecta una sombra cónica idéntica a la producida por la esfera situada sobre la mesa, sólo que vuelta del revés. Los vértices de ambos conos se encuentran en el punto de luz, naturalmente.

En las tres primeras figuras, la sombra de la contrabola se encuentra por encima del plano de la mesa. Sin embargo, cuando la fuente luminosa desciende por debajo de la cima de la pelota, la "contrasombra" cae sobre el plano de la mesa, formando una "contracurva", que es la rama de la derecha de la hipérbola situada bajo la bola. Por así decirlo, ¡el foco desaparecido ha rodeado al infinito y ha vuelto por el otro lado! Como los dos extremos del infinito son uno y solamente un punto, pueden compararse a las extremidades de un anillo cortado y estirado en línea recta. Este anillo infinito y sin mutilación es el trasunto geométrico del famoso pareado de Henry Vaughan: "Contemplé la eternidad la noche pasada. Era como un gran *anillo* de luz pura ilimitada".

Imaginemos la contrabola de la última figura alejándose del punto de luz y expandiéndose continuamente, de forma que siempre sea tangente a todas las generatrices del contracono. Cuando la contrabola llegue a ser lo bastante grande como para tocar la superficie de la mesa, el punto de contacto será el foco de la contrarrama de la hipérbola. Estas dos esferas desiguales, alojadas en los conos, y tangentes al plano secante en los focos de la hipérbola, permiten construir una demostración, clásica y elegante, de que las curvas en cuestión son verdaderamente las ramas de una hipérbola. La demostración se expone con toda claridad en las páginas ocho y nueve de *Geometry and the Imagination*, de David Hilbert y S. Cohn-Vossen (Chelsea Publishing Company, 1952). Si ambas esferas están situadas en el mismo cono, mediante una demostración parecida puede verse que la curva resultante es una elipse (véase el capítulo 15 de mis *Nuevos pasatiempos matemáticos*, Alianza Editorial, Madrid, 1972).

El experimento de las sombras que acabo de describir sirve para poner de manifiesto que las cuatro curvas –circunferencia, elipse, parábola e hipérbola– son secciones planas de un cono. El plano definido por la superficie de la mesa es el plano secante a los conos. Puede observarse que la circunferencia es un caso límite de la elipse. La parábola es caso límite de la elipse y la hipérbola. Lo mismo que ocurre con la circunferencia, las parábolas tienen todas la misma forma, aunque sus tamaños sean diferentes. Sin embargo, existen familias infinitas de elipses e hipérbolas de forma diferente.

Con frecuencia, les resulta difícil a los astrónomos decidir si la trayectoria de un cometa o un meteoro es de tipo elíptico, parabólico o hiperbólico. Es fácil comprender por qué. La más leve perturbación de una parábola la transforma, en un sentido, en una elipse; en el contrario, en una hipérbola. Las trayectorias de los cometas en órbita permanente alrededor del Sol son elípticas. Otros cometas, que llegan a penetrar en el Sistema Solar interior, abandonándolo después para no regresar jamás, describen parábolas o hipérbolas.

Como la hipérbola es una especie de elipse a la que el infinito hubiera partido en dos, no es sorprendente que las dos curvas estén inversamente relacionadas en multitud de formas. Una elipse es el lugar geométrico de todos los puntos de un plano cuya suma de distancias a dos puntos fijos es constante. Estos puntos fijos se llaman focos de la elipse. La propiedad anterior es el fundamento del clásico método de trazado de elipses usando un bucle de hilo que rodea dos alfileres situados en los focos.

La hipérbola es el lugar geométrico de todos los puntos del plano cuyas distancias a dos puntos fijos conserven diferencia constante. La ilustración de la parte superior izquierda de la página 164 muestra un método de trazado de una rama de una hipérbola con un hilo y una regla. El lápiz situado en *P* mantiene el hilo tenso y en contacto con la regla,

mientras ésta gira alrededor de su extremo fijo en el foco A. El hilo está sujeto al foco B y al extremo libre de la regla, C. $BP+PC$ es constante; por tanto, AP-BP tiene que ser constante. Como AP y BP son las distancias desde cada uno de los focos a P, hemos demostrado que la curva es una hipérbola.

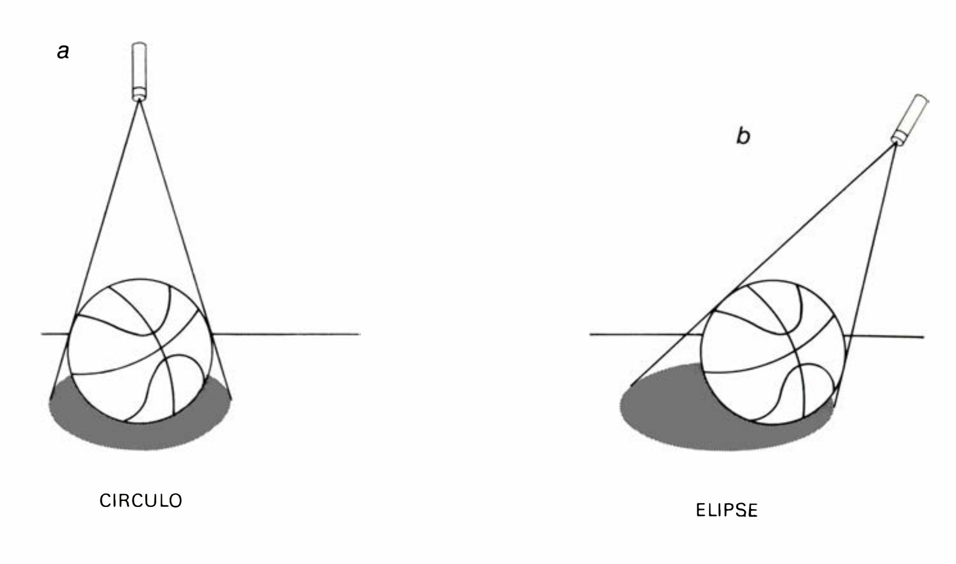
También se obtienen fácilmente elipses e hipérbolas por métodos de papiroflexia, métodos que ponen de relieve la reciprocidad entre ambas curvas. Se dibuja una circunferencia en una hoja de papel translúcido, por ejemplo, papel vegetal. Se marca un punto cualquiera en el interior del círculo, y se pliega el papel muchas veces según diversas líneas, haciendo siempre caer el punto marcado sobre el perímetro de la circunferencia. Cada línea de pliegue es tangente a una elipse; si se hacen suficientes pliegues, la elipse aparecerá como "envolvente" de las tangentes. El punto marcado y el centro del círculo son sus focos. Para obtener una hipérbola basta seguir el mismo procedimiento tomando el punto en el exterior del círculo, en lugar de en su interior. Los focos de la hipérbola vuelven a ser el centro del círculo y el punto elegido.

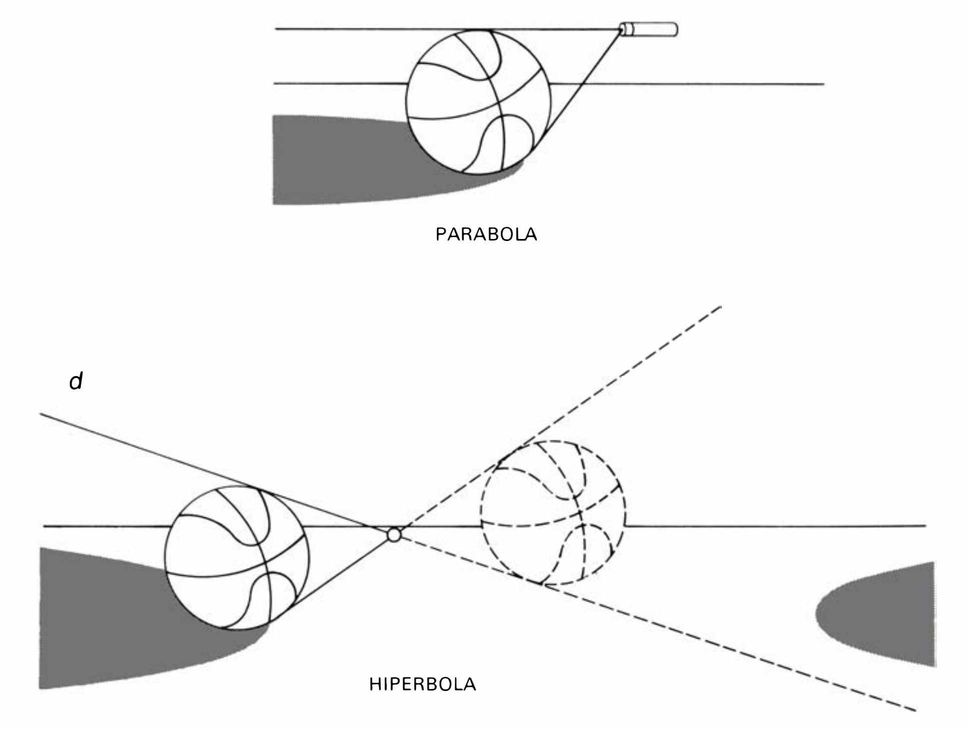
¿Se obtendrá una parábola tomando el punto en la circunferencia del círculo? Desdichadamente, no. El culpable de tan perversa conducta es el foco que la parábola ha perdido. Cada línea de plegado pasa necesariamente por el centro de la circunferencia, siendo, por tanto, tangente a una circunferencia que se ha contraído a un punto. Para obtener una parábola, se requiere un círculo que se haya expandido infinitamente, con lo que su circunferencia se habrá convertido en una línea recta. Tracemos una recta en el papel, y tomemos un punto exterior a ella; el mismo procedimiento de plegado anterior proporcionará una espléndida parábola. El foco que falta es el "centro" del círculo infinito.

De las cuatro secciones cónicas, la menos observable en la vida ordinaria es la hipérbola. Circunferencias y elipses aparecen por doquiera. Se ven parábolas al regar un prado con una manguera o al mirar el vuelo de un balón alto. Una de las escasas maneras de poder observar una hipérbola completa se tiene en la sombra proyectada sobre una pared cercana por una lámpara de pantalla cónica o cilíndrica abierta por ambas bases. Nuestros antepasados pudieron ver una rama de hipérbola observando la sombra proyectada en la pared por una vela de una palmatoria de base circular.

Científicos y matemáticos ven hipérbolas constantemente al representar gráficamente las soluciones de diversas ecuaciones de segundo grado. Incluso la sencilla ecuación $ab=c$, donde c es una constante, se representa gráficamente en una hipérbola. Cientos de leyes físicas tienen esta expresión, entre ellas, por citar sólo dos, las leyes de Boyle y de Ohm, y lo mismo vale para numerosas funciones económicas. Puede realizarse un fácil experimento que mostrará la representación gráfica de $ab=c$ con dos cristales rectangulares planos. Se juntan dos de sus aristas, manteniéndose ligeramente separados los dos lados opuestos mediante una tirita de cartulina o dos cerillas; pueden usarse bandas elásticas para mantener los cristales fijos en esta posición. Introduciendo la base del montaje en agua coloreada, por capilaridad se tendrá la hipérbola que se muestra en la ilustración superior derecha de la página siguiente.

Más adelante, se muestra una hipérbola típica. Las dos rectas de color son las asíntotas de la curva: los límites inalcanzables hacia los que tienden las ramas de la curva al prolongarla indefinidamente. Cuando las asíntotas son mutuamente perpendiculares (aquí no lo son), la hipérbola se llama equilátera o rectangular.

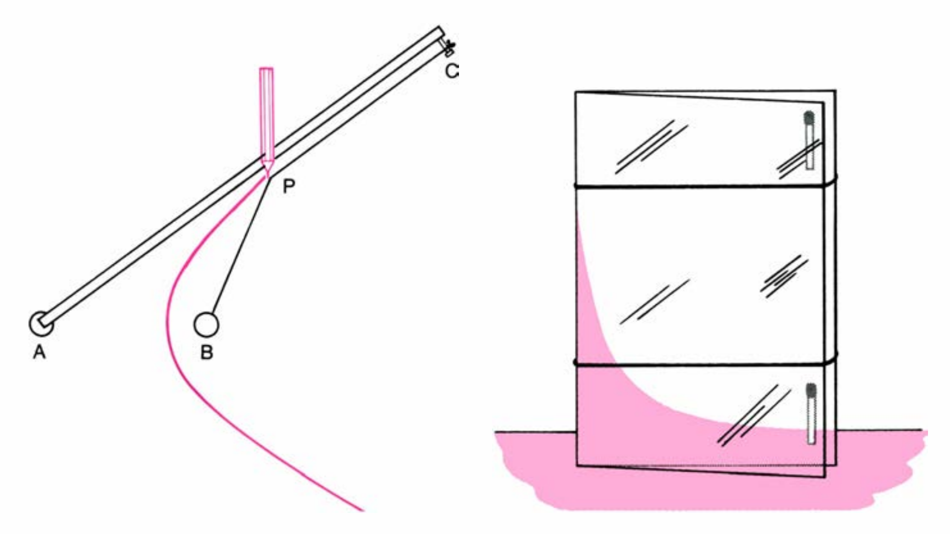




Sombras que muestran las cuatro secciones cónicas curvas

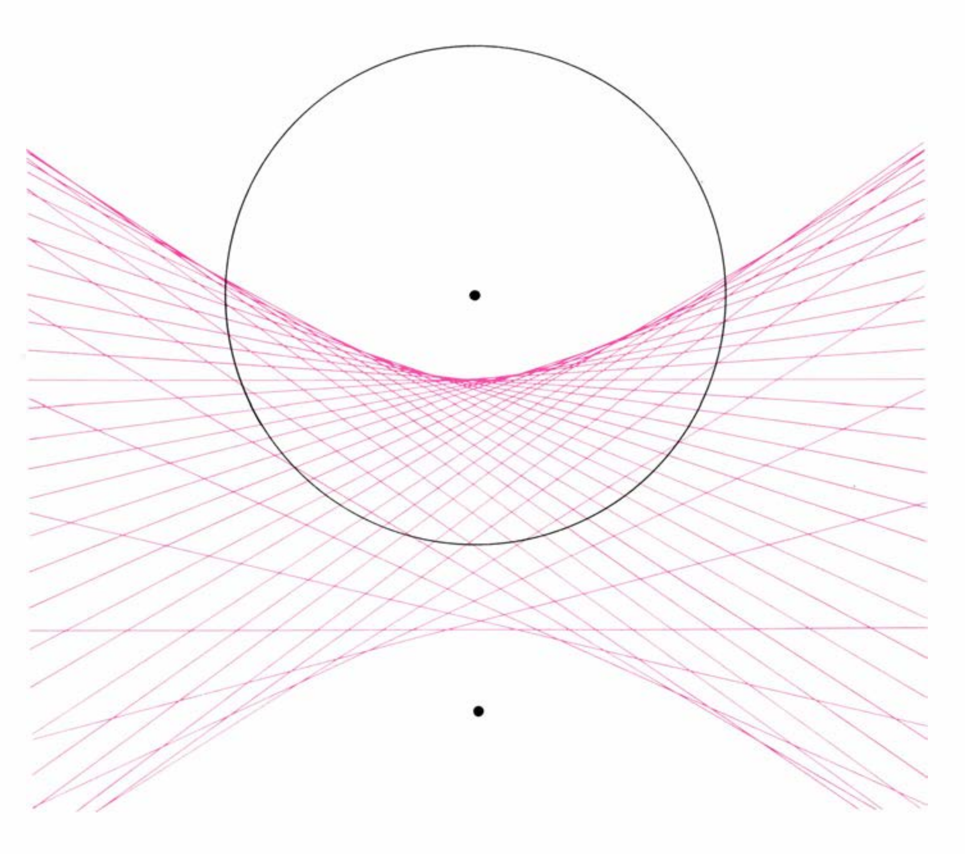
Los brazos de una parábola resultan imposibles de distinguir rápidamente de dos rectas paralelas. En contraste, los brazos de una hipérbola se abren súbitamente en su fuga hacia el infinito, aunque permanecerán confinadas por siempre en el ángulo definido por sus asíntotas. Un filósofo español, Miguel de Unamuno, decía que la hipérbola es una curva trágica. "Creo que si el geómetra hubiera de ser consciente de esta desesperada y desesperanzada lucha de la hipérbola por unirse a sus asíntotas", escribió Unamuno, "¡nos representaría la hipérbola como ser vivo y trágico!"

Sin embargo, en su ensayo "The Inmortality of the Soul", Joseph Addison contemplaba el símil con mayor optimismo. Después de la muerte, el alma se aproxima a Dios, aunque sin llegar a ser Dios. "Todavía no sabemos qué seremos, ni tampoco si llegará a penetrar en el corazón del hombre la idea de la gloria que siempre le estará reservada. El alma, considerada respecto de su Creador, es como una de esas líneas matemáticas que

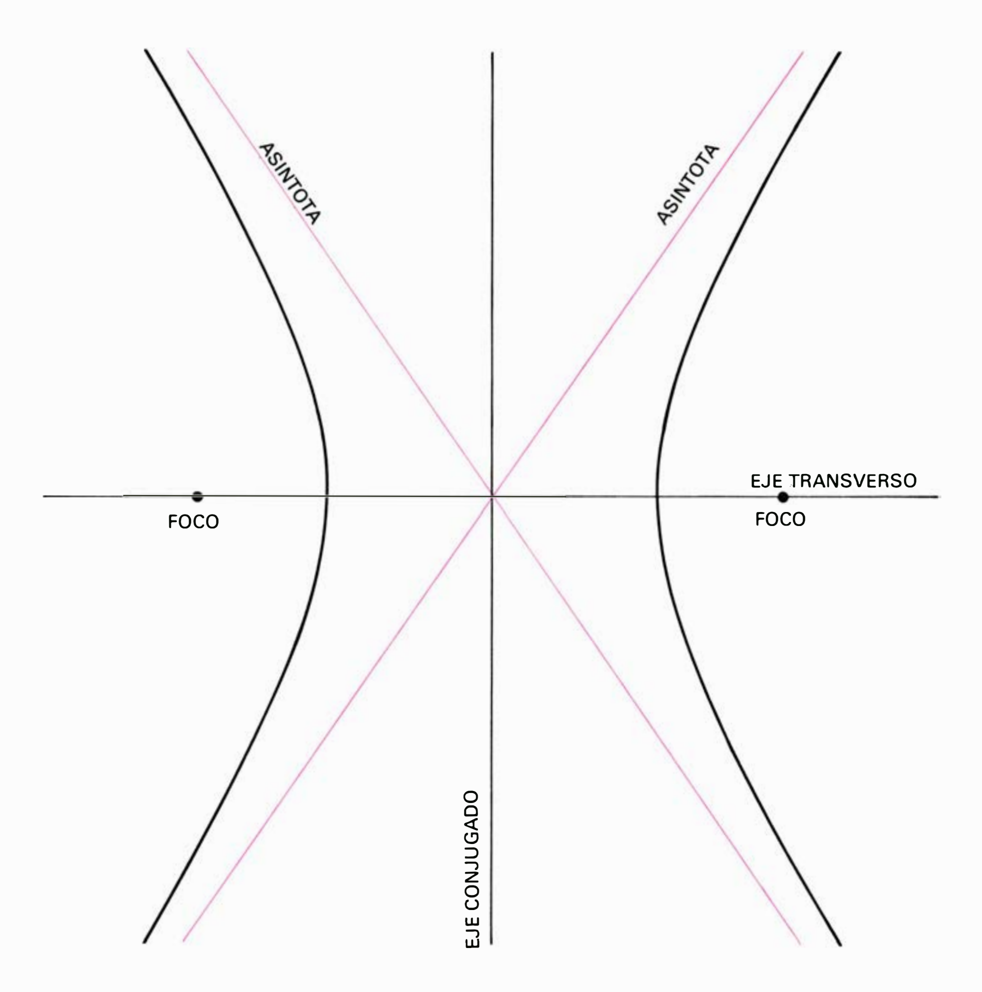


Trazado de una hipérbola con un hilo

Hipérbola producida por capilaridad



Pliegue de una hipérbola



Una hipérbola típica

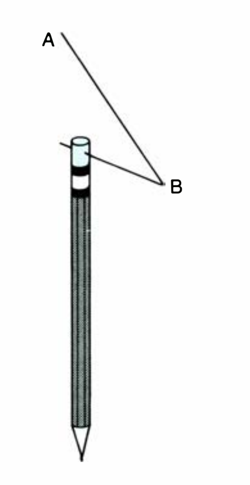
se aproximan indefinidamente durante toda la eternidad, sin posibilidad de llegar a encontrarse jamás."

Las hipérbolas desempeñan un papel fundamental en problemas de determinación de alcances, localización y telemetría. Para ver cómo es esto posible, supongamos que una persona *A* efectúa un disparo de rifle contra un gong lejano, *B*. Admitiendo que el terreno sea plano, ¿dónde deberemos situarnos para escuchar simultáneamente los sonidos procedentes del arma y del gong?

Sea *x* la distancia que el sonido recorre durante el intervalo de tiempo invertido por la bala en viajar desde el rifle hasta el gong. *A* y *B* son focos de un haz de infinitas hipérbolas. Una persona que escuche simultáneamente ambos sonidos tiene que hallarse situada en una rama (la más cercana al blanco) de una hipérbola que sea lugar geométrico de todos los puntos cuya diferencia de distancias a *A* y a *B* sea igual a *x*.

Puede localizarse el lugar donde se ha producido un sonido distante mediante dos pares de puestos de escucha: *A* y *B*, *C* y *D*. Las estaciones *A* y *B* registran independientemente los instantes de recepción del sonido. Sincronizando previamente sus relojes, puede obtenerse con precisión la diferencia entre ambos tiempos. Llamemos *t* a esta diferencia de tiempos. El sonido debe provenir de una rama (la más cercana a la fuente sonora) de una hipérbola que sea lugar geométrico de todos los puntos cuya diferencia de distancias a *A* y a *B* sea igual a *t* veces la velocidad del sonido. Se dibuja esta curva sobre un mapa. Las estaciones *C* y *D* repiten el proceso y trazan la rama de otra hipérbola sobre el mismo mapa. El punto de intersección de ambas ramas más cercano al origen del sonido da la localización de la fuente sonora.

Los sistemas hiperbólicos de navegación, como el LORAN, desarrollado durante la Segunda Guerra Mundial, operan por un procedimiento inverso. En ciertos puntos *A* y *B* de la costa, dos estaciones, llamadas "principal" ("master") y "secundaria" ("slave") emiten señales de radio sincronizadas. Otro par de estaciones principal-secundaria hacen lo mismo desde otros enclaves costeros. Los aviones y barcos en alta mar pueden, midiendo las diferencias de tiempos de recepción de las señales de



Hiperboloide generado por rotación

ambos pares de estaciones, trazar dos hipérbolas, que al cortarse en el mapa permiten determinar la posición.

Pueden encontrarse espejos de sección transversal hiperbólica (acompañados generalmente de espejos de otros tipos) en ciertos telescopios reflectores, en cámaras fotográficas para fines especiales y tras las fuentes luminosas de lámparas de destellos y proyectores. Situando una fuente de luz en un foco de un espejo elíptico todos los rayos reflejados convergen en el otro foco. Si la fuente de luz se encuentra en el foco de una parábola, los rayos reflejados son paralelos, todos ellos dirigidos hacia el foco perdido, situado en el infinito. Un espejo hiperbólico hace que los rayos reflejados diverjan, como se muestra en la ilustración inferior de esta página. Sin embargo, prolongando hacia atrás estos rayos divergentes, en la forma indicada en línea de trazos, vuelven a reunirse cordialmente en el otro foco. En cierto sentido tales rayos han atravesado el infinito y encontrado el foco perdido en un punto situado detrás de sus orígenes.

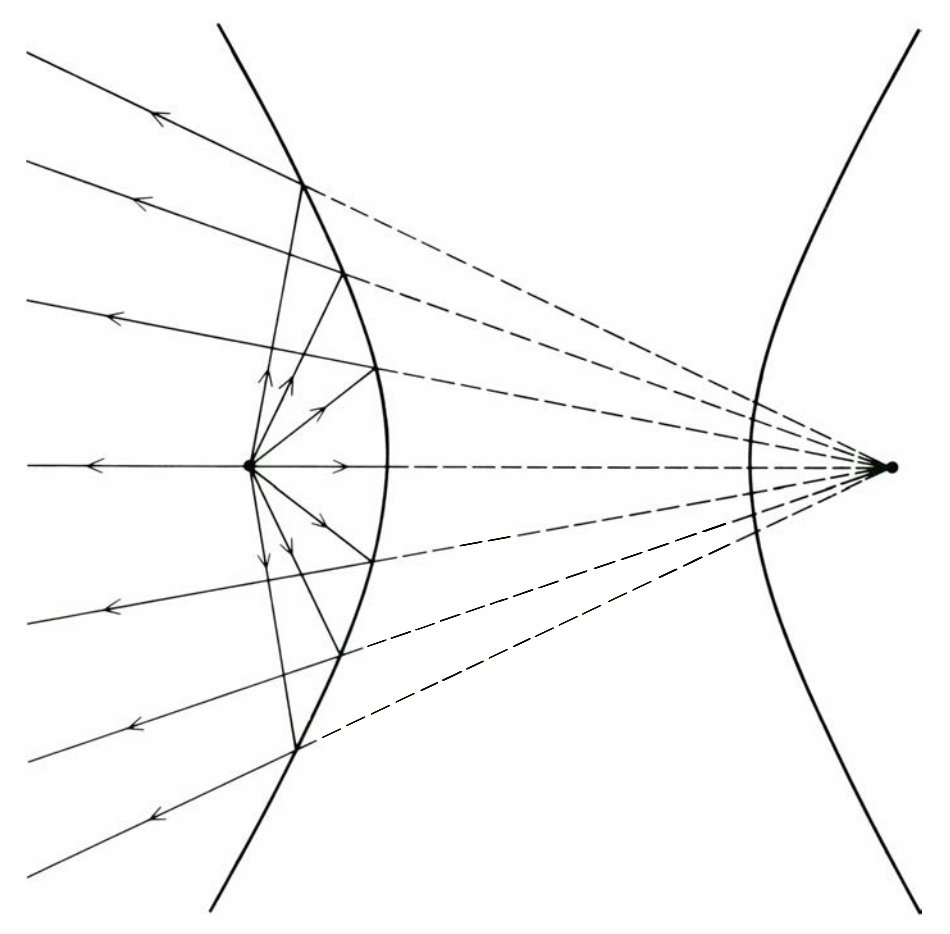
El "hiperboloide de una hoja" es una preciosa superficie que posee notables propiedades; fue ya Arquímedes el primero en describirla. Se muestra un modelo de hilos en la ilustración superior (centro) de la página siguiente. Las secciones transversales verticales son hipérbolas, y las horizontales, elipses. Cuando las secciones transversales son circunferencias, es un "hiperboloide de revolución de una hoja", así llamado porque puede generarse haciendo girar una hipérbola alrededor de su eje conjugado. (Si se hace girar la hipérbola alrededor de su eje transversal, se genera un hiperboloide de revolución de dos hojas: un par de estructuras en domo separadas una distancia finita.)

En 1669, Christopher Wren, arquitecto que diseñó la catedral de San Pablo, comunicó un extraordinario descubrimiento relativo al hiperboloide de una hoja: pudo demostrar que el hiperboloide de una hoja es lo que ahora se llama una superficie reglada, ¡una superficie formada por una infinidad no numerable de líneas absolutamente rectas!

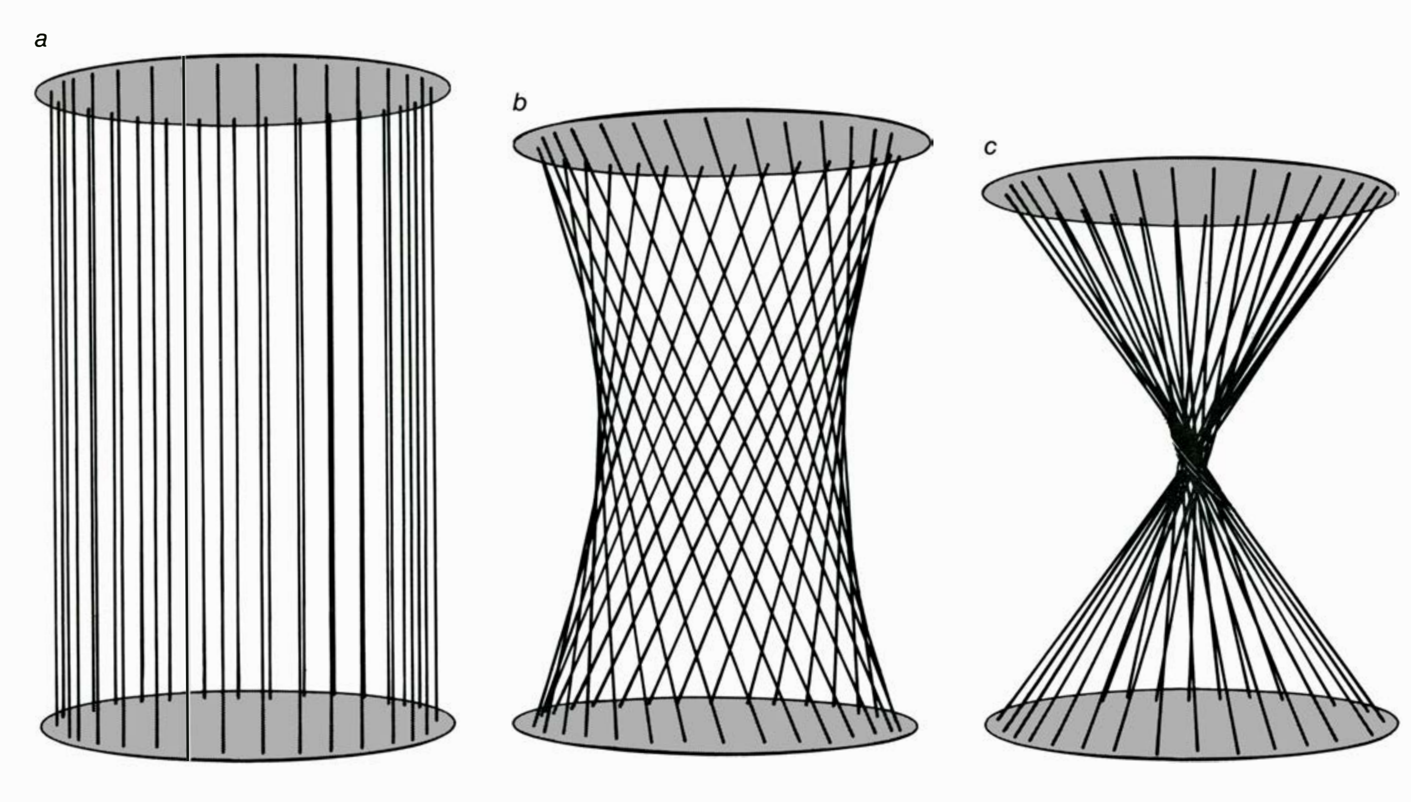
Así, por ejemplo, un cilindro es una superficie reglada formada por líneas rectas paralelas. Un cono es una superficie reglada formada por líneas rectas que se cortan en el vértice del cono. El hiperboloide es una superficie reglada. formada por dos familias distintas de rectas. En el centro de la ilustración superior de la página opuesta pueden verse algunas de ellas, pertenecientes a uno de tales conjuntos, todas ellas inclinadas en igual manera, y sin que se corte ningún par de dichas rectas. La otra familia (que no se muestra en la figura) es imagen especular de la anterior. Por cada punto del hiperboloide pasa una línea recta de cada una de estas familias. El par de rectas que pasa por cada punto define el plano tangente en él.

Observando el modelo de hilos es fácil comprender que el hiperboloide de una hoja puede generarse por rotación alrededor de un eje de un segmento rectilíneo oblicuo y no coplanario con él; es decir, que la recta que lo contiene no sea paralela al eje ni tampoco lo corte. (Si el segmento en rotación es paralelo al eje, genera un cilindro; y si al prolongarlo corta al eje, genera un tronco de cono.) Esta propiedad sugiere un sencillo experimento a realizar con un lápiz y un clip sujetapapeles. Se abre el clip de modo que forme un ángulo agudo, y se clava uno de sus extremos en la goma de borrar del lápiz. Se retuerce el alambre con el fin de que la parte vertical *AB* quede oblicua con el eje vertical del lápiz. Colocando el lápiz entre ambas manos, se lo hace girar rápidamente deslizando las manos adelante y atrás. Si la iluminación es adecuada, el segmento oblicuo formará un hiperboloide transparente.

Si se hace girar rápidamente un cubo sobre uno de sus vértices, sus seis aristas oblicuas generan una superficie similar. Con un poco de práctica se puede aprender a sujetar el dado entre el índice y el pulgar, y, chasqueando los dedos, hacer girar el dado como si fuera una peonza sobre uno de sus vértices. Bajando la cabeza hasta ver el cubo giratorio desde un costado, se observará el perfil de un hiperboloide entre dos conos.



Los rayos procedentes de un foco al reflejarse en la hipérbola se reúnen en el otro foco



Al retorcer el modelo cilíndrico (**izquierda**) *se obtiene un hiperboloide* (**centro**) *y un doble cono* (**derecha**)

No es difícil construir el modelo de hilos de la ilustración superior de esta misma página. Basta pasar un hilo adelante y atrás por una serie de orificios igualmente separados a lo largo de dos discos de cartón o madera contrachapeada. Para evitar que los hilos se deslicen, es suficiente pegarlos con cola a los orificios. Si los discos se separan lo más posible, manteniendo todos los hilos verticales, se obtiene un modelo de cilindro, como en la parte izquierda de esa misma ilustración.

Si uno de los discos se gira 180 grados en sentido horario, los hilos configuran el modelo de dos conos. Entre ambos casos límite se puede obtener una familia infinita de hiperboloides, dependiendo de la amplitud del giro realizado, reglados con hilos rectos sesgados por igual en un mismo sentido. Al girar el disco en sentido antihorario se obtiene el conjunto de todos los hiperboloides simétricos de los anteriores respecto a un plano, reglados por hilos rectos sesgados en sentido contrario.

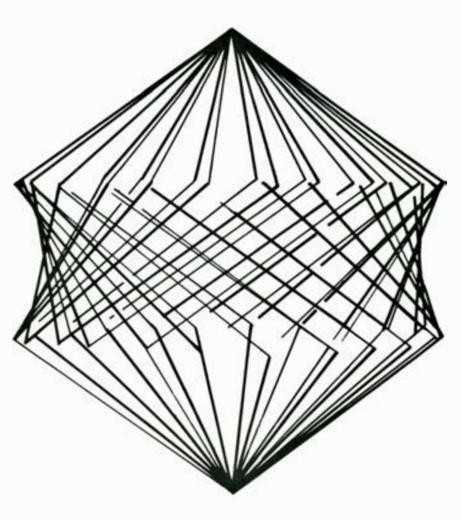
En las páginas 16 y 17 de *Geometry and the Imagination* se describe un modelo mucho más difícil de construir, realizado con alambres rígidos en vez de hilos. El par de alambres de cada intersección está unido mediante una junta universal, que permite giros pero no deslizamientos. Podría esperarse que tal estructura fuese rígida; en cambio, es flexible de modo curioso. Comprimiendo el modelo en una dirección, las secciones transversales elípticas degeneran en segmentos rectilíneos, y las varillas se pliegan hasta formar un haz plano de segmentos rectilíneos cuya envolvente es una hipérbola. Aplastándolo en la otra dirección, las varillas forman en el plano horizontal un haz cuya envolvente es una elipse. La ilustración de la página siguiente, basada en una fotografía de *Geometry and the Imagination,* muestra en qué forma dos hiperboloides de revolución pueden producir una transmisión de engranajes entre ejes oblicuos no situados en un mismo plano. Los dientes de los engranajes son uno de sus conjuntos de generatrices rectilíneas.

Se tiene un notable ejemplo de aplicación arquitectónica de un hiperboloide de revolución de una hoja en el Planetario McDonnell de Forest Park, en San Luis. El proyectista Gyo Obata eligió esta superficie porque las trayectorias hiperbólicas de ciertos cometas hacen sentir –en sus propias palabras– "el dramatismo y la excitación de la exploración del espacio". Obsérvese la línea recta definida por la sombra proyectada por el techo circular del planetarium. ¿Es esta línea de sombra una de las rectas generatrices de la superficie, o es una curva del espacio que solamente aparece como recta al mirarla desde el ángulo indicado? Se dará la solución el mes próximo.

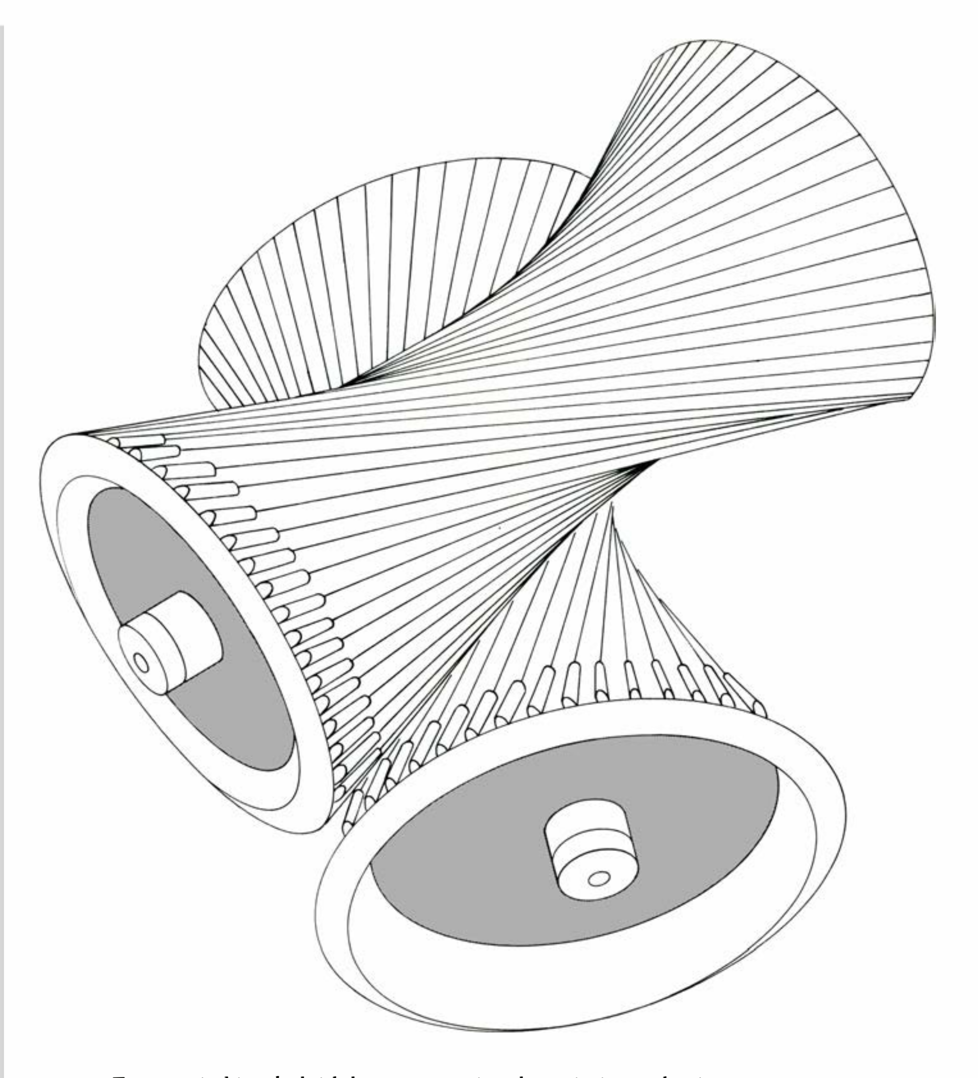
El texto cifrado del mes pasado –el que Edgar Allan Poe no pudo resolver– tiene el siguiente texto llano:

"Mr. Alexander,

How is it that the messenger arrives here at the same time with the Saturday courier and otlfer Saturday papers when according to the date it is published three days previous? Is the fault with you or the postmasters?"



Al girar un cubo se ve un hiperboloide entre dos conos



Engranaje hiperboloidal que transmite el movimiento de ejes que se cruzan

(D. Alejandro,

¿A qué se debe que el mensajero llegue aquí al mismo tiempo que el correo del sábado y demás prensa del sábado, si de acuerdo con su fecha se publica tres días antes? ¿Es culpa suya o del servicio de correos?).

La clave consiste en una sustitución polialfabética que utiliza doce alfabetos cifrados por las palabras "United States". Cada letra de ellas indica el grado de desplazamiento de una clave cesárea. Así el alfabeto-clave para M, la primera letra del texto llano, es A=U, B=V, C=W,... y así sucesivamente. Para la R, que es la segunda letra del texto llano, la clave es A=N, B=O, C=P, etcétera.

El criptograma publicado contenía 16 errores: en primer lugar, la tercera letra de UNITED es una I, que el autor de la clave tomó como una J; en segundo lugar se omitió la quinta letra del mensaje. Si no se hubiera cometido este segundo error, Poe habría adivinado que el encabezamiento era "Mr. Alexander", y la solución completa no le hubiera resultado costosa. El próximo número de *Cryptologia* (véase la sección de "Juegos matemáticos" de INVESTIGACION Y CIENCIA del mes de octubre) explicará cómo puede analizarse por métodos modernos las claves de este tipo.

En el número de julio se presentó en esta sección el paradójico juego siguiente. En cada una de las fichas de una colección infinita están escritos dos enteros consecutivos. Hay 10^n copias de cada ficha, siendo n el menor de los dos números que constan en ella, y $n = 1,2,3,4,...$ Un árbitro selecciona al azar una de las tarjetas, y después escribe uno de los números en la frente de uno de los jugadores, y el otro número, en la del otro. Los jugadores acuerdan que el poseedor del número más pequeño, k, ha de pagar k dólares al otro. El jugador A, que ve en la frente de su oponente el número m razona así: Mi número es o bien $m+1$ o bien $m-1$. Dado que hay diez veces más fichas que contengan los números m y $m+1$ que fichas que tengan $m-1$ y m, mis probabilidades de ganar son 10 veces

mayores que las de perder. Evidentemente, B razonaría del mismo modo. ¿Cómo es que ambos jugadores salen favorecidos en este juego?

Más de 50 lectores señalaron inmediatamente que la paradoja proviene de la falsa hipótesis de que los números pueden elegirse al azar del conjunto infinito de enteros positivos con la misma probabilidad distinta de cero. Puede ponerse de manifiesto lo absurdo del razonamiento de multitud de formas. Tomemos un número k cualquiera. La probabilidad de que un entero positivo seleccionado "al azar" sea menor o igual que k es nula; y la probabilidad de que sea mayor que k es igual a 1. Dicho de otra forma, hay una probabilidad nula de elegir un número lo suficientemente pequeño como para poder escribirlo en una superficie finita con símbolos de tamaño finito, por pequeño que éste sea.

George Peter Wacktell puso de manifiesto la idea central de la siguiente forma: No se puede barajar un mazo infinito de cartas marcadas con todos los enteros positivos, porque si se pudiera resultaría el siguiente teorema contradictorio: "Para todo par de cartas tomadas de tal mazo barajado es infinitamente probable que cada una de ellas tenga un número superior al de la otra". Dicho brevemente, tal juego no puede llevarse a cabo.

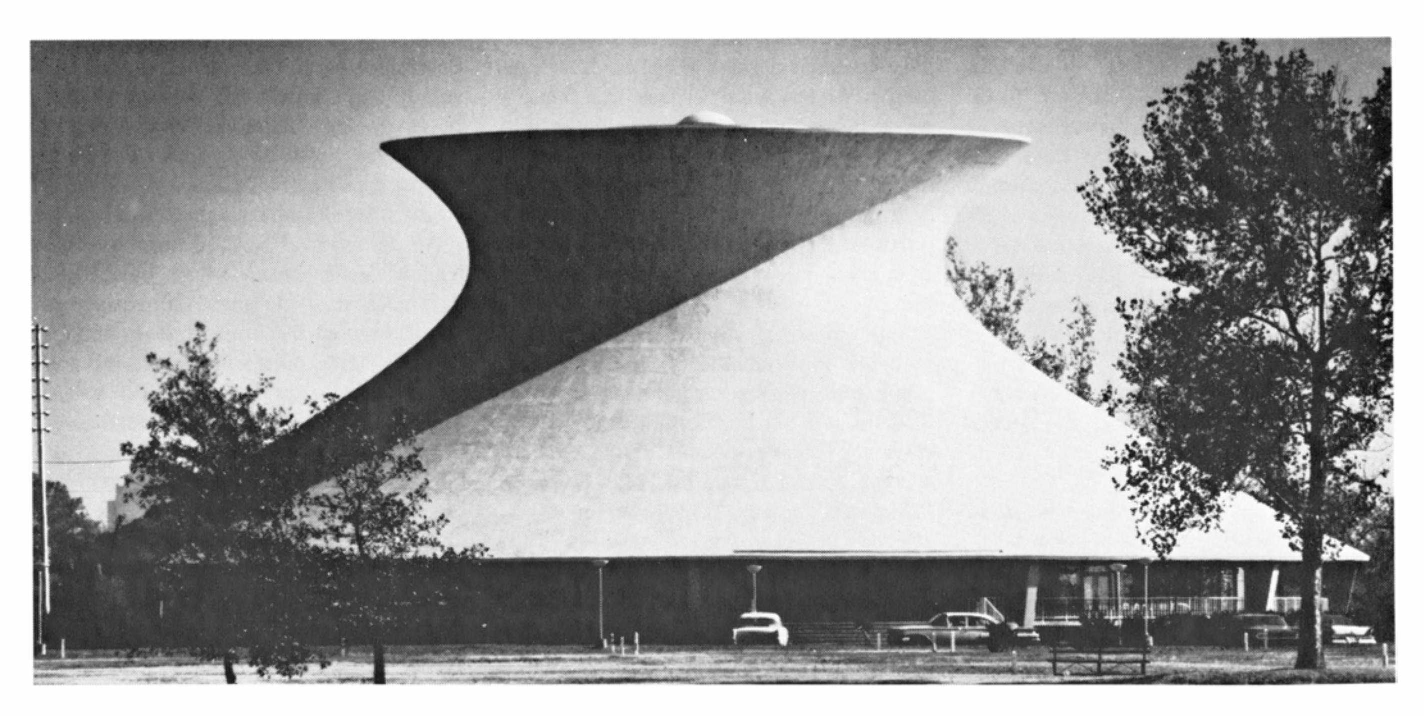
Entre los lectores que enviaron análisis especialmente interesantes por su detalle y profundidad se encuentran Steven J. Brams, Elkan Halpern y Thomas Louis, Roger B. Lazarus, Frederick Mosteller y Herbert Robbins.

Supongamos que el mazo de cartas sea finito, con lo que el árbitro podrá realmente barajarlos al azar y elegir un par de números lo bastante pequeños como para poder escribirlos. No se conoce la cota superior de los pares de números. Al menos un jugador se negará a jugar en todos los casos. I. Martin Isaacs, matemático de la Universidad de Wisconsin, extendió y corrigió anteriores observaciones mías sobre este juego, en la siguiente forma:

"Si las cartas están numeradas (1,2), (2,3), (3,4),... (donde posiblemente falten algunas cartas y otras estén repetidas, o incluso en el caso de que haya infinitas cartas, pero el procedimiento de selección tenga una distribución probabilística no uniforme), todo jugador que vea en su oponente un número par debe negarse a jugar. Si ve un número impar, no tiene necesidad de imponer su veto ni aún (sorprendentemente) en el caso de que supiera que el número impar que está viendo es el más alto de todo el mazo. Esto se encuentra en contradicción con su análisis, que exige imponer su veto al que vea el número mayor, incluso si el número es impar. Tal jugador puede consentirse jugar, confiando en la racionalidad del otro jugador, quien al ver un número par no permitirá que llegue a producirse el pago.

"Para demostrar mi afirmación, observemos que un jugador que vea un 1 no puede perder, por lo que no necesita interponer el veto. Así pues, es juego racional vetar en todos los números mayores que 1, y jugar al ver 1. (Evidentemente, no es ésta la única forma racional de jugar.) Cada jugador, por ser racional, ha de precaverse contra todas las estrategias racionales de su oponente. Así que, si un jugador ve un 2, ha de vetar el juego para salvaguardarse contra la estrategia racional 'sí al 1, no a todos los demás'. Resulta ahora que también es racional 'sí al 3, no a todos los demás', porque si un jugador ve un 3 se sabe seguro, dado que si tuviera un 2, su oponente ha de vetar el juego al ver un 2. Como protección contra la estrategia 'sólo al 3', un jugador racional ha de vetar un 4, con lo que 'sí tan sólo al 5' resulta ser estrategia racional. Siguiendo en esta forma, se demuestra mi aserto por inducción.

"Ocurre aquí algo muy curioso: El análisis anterior subsiste por completo si falta del mazo la carta (1,2). Sin embargo, si el árbitro anunciase que faltaba la carta (1,2), el juego es equivalente al anterior, con la excepción de que ahora deben vetarse los números impares. Por otra parte, si cada jugador repasa secretamente el mazo y comprueba que falta (1,2), pero no sabe que su oponente tiene también esta información, la exigencia de racionalidad requiere vetar los pares (excepto, claro está, al 2). Se trata otra vez de una cuestión del tipo 'A sabe que B sabe que...' En toda la discusión anterior estoy suponiendo que cada jugador impone su veto mediante voto secreto."



El Planetarium McDonnell, en San Luis

Taller y laboratorio

El agua caliente se congela más aprisa que el agua fría. ¿Cómo es ello posible?

Jearl Walker

A veces se oye decir que el agua caliente se congela antes que el agua fría. La creencia es bastante antigua; las gentes de regiones frías, como Canadá, lavan sus coches o llenan las fuentecillas donde beben los pájaros en su jardín con agua fría, creyendo que con ello retrasarán la congelación. ¿Es cierto esto? La respuesta es afirmativa, aunque el resultado, que va en contra de lo que intuitivamente se diría, carecía de base hasta que Erasmo Mpemba, estudiante de la escuela superior de Tanzania, redescubriera el efecto mientras elaboraba helado (de crema) como parte de un trabajo práctico de física.

El experimento consistía en calentar leche, echarle azúcar, mezclar, enfriar la mezcla hasta temperatura ambiente y después depositarla en el congelador. Cierto día, Mpemba y un compañero suyo colocaron sus respectivas mezclas en el frigorífico al mismo tiempo, pero debido a que los dos tenían prisa, ni Mpemba se preocupó de enfriar su mezcla ni su compañero en calentar la suya. Mpemba se sorprendió cuando vio que su mezcla caliente se había congelado mucho antes que la mezcla tibia.

El efecto del helado ganó en credibilidad y fama cuando Mpemba y D. G. Osborne, del Colegio Universitario de Dar es Salaam, publicaron un trabajo sobre el mismo en "Physics Education". Este trabajo despertó una fecunda polémica, tanto en ésa como en otras publicaciones. Los científicos que escribieron sobre el trabajo mantenían, unos, que el efecto era ya suficientemente conocido y, otros, que la forma de hacer el experimento había sido errónea. El efecto es real, indiscutiblemente, pero la explicación del hecho sigue siendo motivo de controversia.

Una de las primeras explicaciones que se dieron fue que el recipiente más caliente hacía fundir el hielo subyacente, produciéndose así un mejor contacto térmico con la estantería del congelador. Sin embargo, el efecto es el mismo si el recipiente está térmicamente aislado de la estantería del congelador, como en el caso del experimento de Mpemba y Osborne. Además, esta explicación no es válida para los ejemplos citados, como lavar coches y llenar los bebederos de los pájaros, ni tampoco lo es cuando se usan frigoríficos de descongelación automática, en los que no se forma escarcha.

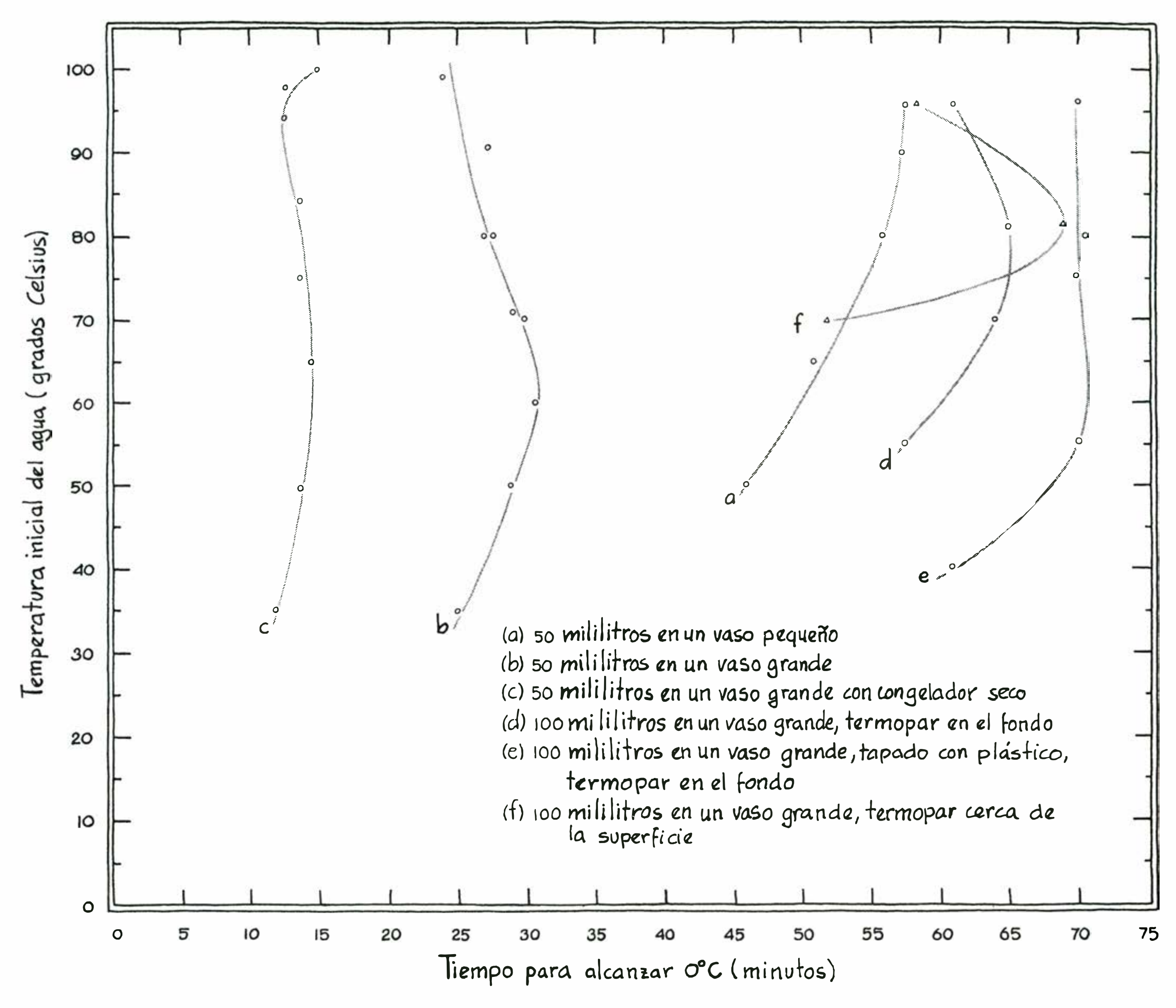
Al parecer, hay varios factores que están implicados en el efecto. En primer lugar, en el recipiente más caliente el líquido puede circular mejor; de esta manera, el agua caliente de la región central se mueve con rapidez hacia las paredes del recipiente o hacia la superficie superior del agua. En segundo lugar, si el agua es más caliente, se puede liberar una mayor cantidad de gas disuelto en el agua. Los gases disueltos retrasan el enfriamiento y, por tanto, su eliminación antes de enfriarse permite que el agua alcance más pronto el punto de congelación. Por ejemplo, es mucho más probable que se congelen las cañerías de agua caliente de una casa que las de agua fría, lo cual se debe a que los gases disueltos en el agua se eliminan en presencia del calor. En tercer lugar, el agua más caliente quizá pierda mayor cantidad de masa y calor al evaporarse que el agua fría. Por ello, habría que enfriar menos masa y se llegaría antes al punto de congelación. Si hay una pérdida significativa de masa, entonces (una vez alcanzado el punto de congelación) el agua que inicialmente estaba más caliente se congelará antes, ya que es menor la masa a partir de la cual se debe liberar el calor para conseguir el paso de agua líquida a hielo.

Voy a describir aquí mis experimentos sobre la congelación del agua y sugeriré otros que se pueden realizar en casa; de cualquier forma, no he sido capaz de resolver la controversia. Eso se lo reservo al lector. Como decía Ian Firth, de la Universidad de St. Andrews, en uno de sus más profundos estudios sobre el tema: "Hay una gran cantidad de variables de experimentación en el problema, de manera que se puede garantizar que los resultados de llevar a cabo el experimento en un laboratorio serán diferentes de los que se realicen en cualquier otro". La primera dificultad con que se encontrará el lector en sus experimentos de congelación será la de estudiar sólo una variable cada vez.

Con el fin de tipificar la clase de recipiente, hice los experimentos con vasos graduados de pirex ordinario. El pirex es material adecuado para este experimento, porque resiste los cambios bruscos de temperatura desde la estufa hasta el congelador, sin romperse. Se colocó una sonda de un termopar en un vaso lleno de agua por debajo del nivel de la misma, y se fijó en una determinada posición. La otra sonda del termopar se introdujo en un baño de agua helada para tener una temperatura de referencia, ya que la temperatura de dicho baño es la de cero grados Celsius. Cualquier cacharro de cocina sirve, a buen seguro, para contener el baño de referencia, aunque lo más conveniente es utilizar un termo. En su interior, el baño de agua helada dura varias horas. Si no quiere tener problemas con el manejo del termopar, puede servirle un termómetro común donde se pueda leer desde 0 hasta 100 °C. Asegúrese de que el depósito de mercurio del termómetro esté totalmente sumergido en el agua.

Para aislar térmicamente el vaso de la escarcha formada sobre la repisa del congelador, usé un tapete de corcho de los que se suelen colocar debajo de los platos calientes. Se obtiene el mismo resultado utilizando poliestireno o cualquier trapo o paño de tela gruesa. Si no aisláramos el recipiente, nos encontraríamos con que el agua más caliente se enfría más rápido, en parte porque funde más rápidamente la escarcha que la rodea. Este aislamiento no es necesario si se trata de un frigorífico de descongelación automática. En mi congelador, la temperatura del aire se situaba entre –8 y –15 grados Celsius. Procure mantener cerrada herméticamente la puerta de su congelador para conservar, en su interior, una temperatura estable.

Al calentar el vaso de precipitados, hágase lentamente, sobre un quemador eléctrico o de gas. A fin de asegurar la calefacción uniforme de toda la masa de agua, ponga una placa metálica plana sobre el quemador; de lo contrario, puede ocurrir que el par termoeléctrico se encuentre situado encima de una porción del quemador particularmente caliente o más fría y dé lugar a lecturas de temperatura erróneas. Se debe tapar el vaso para que el agua evaporada durante la calefacción vuelva a caer en su interior. Lo más aconsejable es trabajar primero con una pequeña cantidad de agua, pongamos unos 50 o 100 mililitros, a fin de reducir el tiempo necesario para la obtención de datos a partir de varias



Experimentos de congelación de agua partiendo desde diferentes temperaturas iniciales

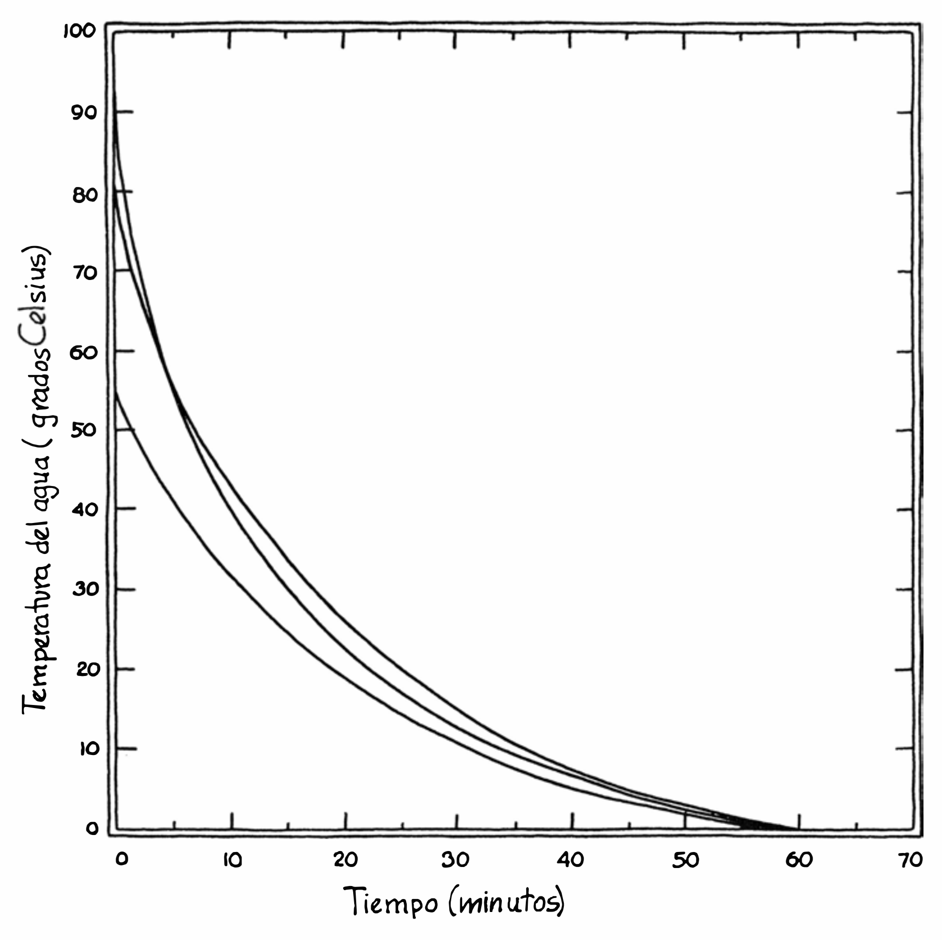
temperaturas iniciales distintas. El factor tiempo también es uno de los motivos por los que se debería practicar varias veces el proceso antes de acometer realmente la toma minuciosa de datos. Yo empecé trabajando con 200 mililitros de agua, tomé datos durante tres días y luego me encontré con que no me había esmerado lo suficiente en el procedimiento. Tuve que desechar todos los datos.

Cuando el agua que está sobre el hornillo alcance la temperatura inicial deseada, se pasa al congelador, junto con la sonda del par termoeléctrico que lleva acoplado; hay que actuar con rapidez porque, en los primeros minutos después de quitar el vaso de la fuente de calor, el agua se enfría en seguida. Si el agua se calienta previamente en una tetera y luego se vierte en el interior del vaso que ya se encuentra en el congelador, tomando entonces una lectura de temperatura, es imposible obtener lecturas correctas. Para entonces, el agua ya se ha enfriado demasiado.

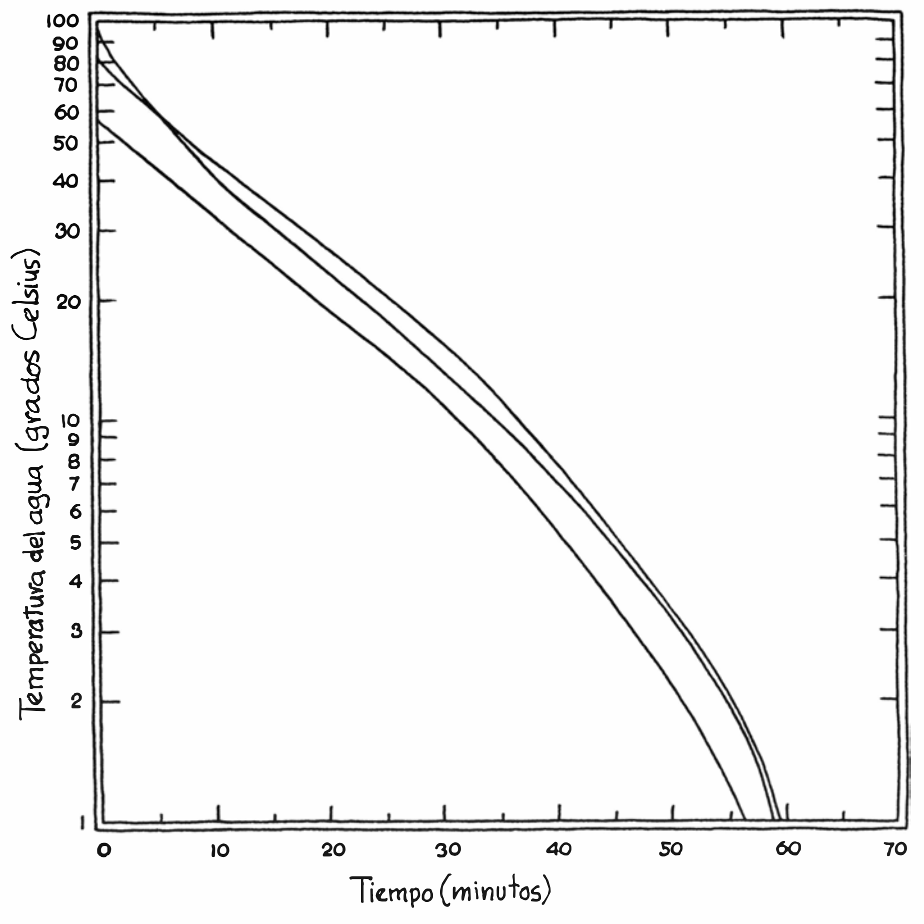
Yo trabajé con agua corriente del grifo; la hervía para eliminar el gas disuelto. Después de cada operación, dejaba que el agua volviera a la temperatura ambiental, antes de añadirle más agua para restablecer el volumen original. La medida de la cantidad de agua necesaria para restablecer el volumen original en el interior del vaso nos da una medida de la cantidad de agua perdida por evaporación durante el proceso de enfriamiento. Comprobé que la manera más precisa y práctica para añadir el agua era inyectarla con una jeringa hipodérmica marcada con una escala en centímetros cúbicos.

La sonda del par termoeléctrico se colocó a distintas profundidades en el interior del agua, a fin de poder controlar la velocidad de enfriamiento a varios niveles de profundidad. Hay que evitar que la sonda quede expuesta al aire durante la operación debido a la caída del nivel del agua por evaporación, porque entonces la sonda sólo mediría temperatura del aire. Si se dispone de un frigorífico de descongelación automática y de un ventilador que se conecte y desconecte automáticamente, es muy posible que se obtenga una oscilación sospechosa en las lecturas de temperatura. En este caso, hay que intentar mantener el ventilador o bien conectado o bien desconectado durante la totalidad del proceso. La corriente de aire creada por el ventilador es un factor significativo en la refrigeración del vaso de precipitados.

En lugar de medir el tiempo necesario para congelar toda la masa del agua, detenía el proceso cuando el par termoeléctrico alcanzaba cero grados Celsius. Para entonces, solía haberse formado una fina capa de hielo en el borde exterior de la superficie del agua. También se podría medir el tiempo hasta la conge-



Datos tomados cerca del fondo en el recipiente que contenía 100 mililitros de agua



Los mismos datos representados en escala semilogarítmica

lación completa, colocando el par termoeléctrico en el centro de la muestra y esperando a que la temperatura empiece a descender por debajo de cero grados Celsius. La temperatura sólo podrá estar por debajo de cero grados cuando el agua se halle totalmente helada.

La mejor forma de comprobar la fiabilidad del procedimiento que se ha seguido es que pueda reproducirse. Sin embargo, por muy reproducibles que fueran mis mediciones, seguía obteniendo grandes desviaciones en alguna operación, lo cual era muy extraño. ¿A qué se debía? No lo sé, pero sospecho que las impurezas del agua alteraban el tiempo necesario para enfriarla.

El primer experimento lo hice con 50 mililitros de agua en un vaso pequeño, graduado hasta 150 mililitros, con un diámetro interno de 5,3 centímetros y una altura de 8,1 centímetros. Utilicé un refrigerador con un congelador no autodescongelable. Una representación gráfica del tiempo necesario para enfriar el agua hasta cero grados Celsius frente a la temperatura inicial del agua presentaba un incremento uniforme hasta los 80 grados de temperatura inicial. Luego, el aumento de la curva de enfriamiento se hacía menos marcado. Lo más decepcionante era que la curva no presentaba una caída en su tramo final, caída que demostraría que el tiempo necesario para una temperatura inicial alta es menor que para una temperatura más baja.

En cambio, al utilizar la misma cantidad de agua en un vaso mayor (graduado hasta 600 mililitros, 8,2 centímetros de diámetro, 12,4 centímetros de altura y la columna de agua alcanzando 0,9 centímetros) y en el mismo congelador, la curva obtenida sí presentaba un codo correspondiente a temperaturas iniciales entre 60 y 65 grados Celsius. Este resultado significa que el agua que inicialmente está a unos 99 grados alcanza el punto de congelación aproximadamente siete minutos antes que el agua que inicialmente está a 60 grados. Es muy probable que, si yo hubiera medido la longitud del proceso hasta la total congelación del agua, el codo de la curva fuera todavía más pronunciado en el supuesto de que la evaporación del agua constituya un factor significativo. A temperaturas más elevadas, la evaporación sería mayor y se perdería más masa, de tal modo que sería menor el calor que debería extraerse del agua a cero grados Celsius para transformarla en hielo.

Los resultados sugieren que la evaporación es importante. Al pasar de un vaso pequeño a otro mayor, la superficie del agua alcanzó una extensión superior en más de dos veces la anterior, y,

por tanto, aumentó cuantitativamente la evaporación. Lo que no sé es de qué modo el acortamiento de la columna de agua afecta a la trayectoria de las corrientes de la convección en el interior del agua. Puede ser que la relación entre el área de la superficie y el área de los lados constituya un valor cuantitativo más significativo para describir el cambio de un vaso a otro. En la superficie del agua, la pérdida de calor se debe principalmente a la evaporación, mientras que por los lados se pierde por convección y radiación. La relación de áreas mencionada aumentó tres veces al pasar de un vaso a otro mayor. Así pues, la evaporación tomó ventaja respecto a la pérdida de calor a través de las paredes del vaso, y quizá sea por este motivo por lo que la curva de tiempo presentaba un codo más cerrado.

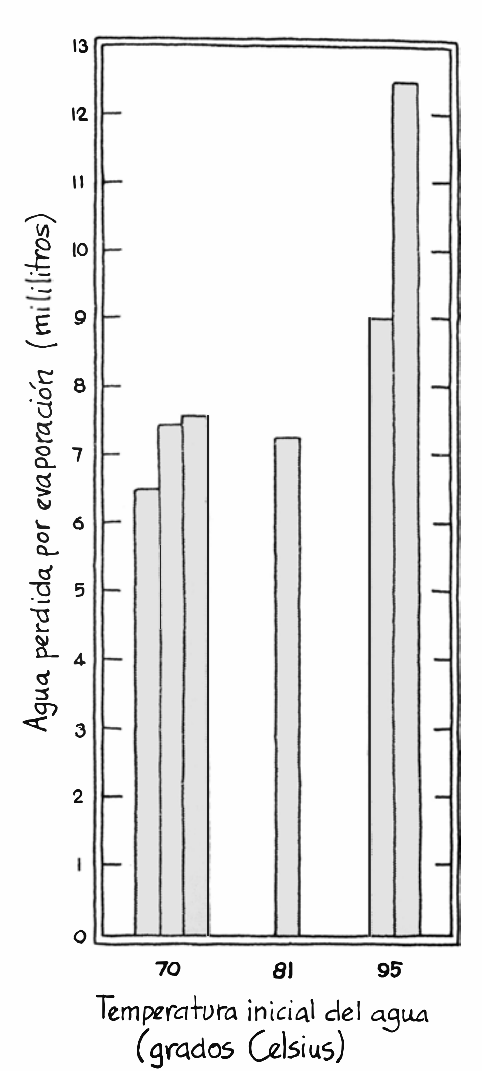
Tomando la misma cantidad de agua en el mismo vaso pero utilizando un frigorífico de descongelación automática, el agua se refrigera en la mitad de tiempo aproximadamente, puesto que hay una corriente de aire constante alrededor del recipiente. En este caso también se puede observar el codo de la curva, que aparece a la misma temperatura aproximadamente. A temperaturas iniciales próximas al punto de ebullición, parece que la curva vuelve a subir. Quizás en este punto el aumento de temperatura empiece a predominar sobre la pérdida de masa debida a la evaporación o sobre el incremento de la circulación del agua. Es posible que usted decida centrar su estudio en este segundo incremento y obtener sus propios datos al respecto; de momento, no hay nada publicado acerca de ello.

Posteriormente, pasé a aumentar el volumen de agua a 100 mililitros, utilizando el vaso de precipitados de 600 mililitros y el congelador inicial (no seco). Tomé datos con la sonda del par termoeléctrico cerca del fondo y luego cerca de la superficie. El fondo dio lugar a una curva con un codo bastante marcado, pero el pico se corrió hacia una temperatura inicial mayor, de aproximadamente unos 80 grados Celsius. El agua que inicialmente se encontraba a 95 grados alcanzaba el punto de congelación unos tres minutos antes que el agua que inicialmente estaba a 80 grados.

Para observar en qué momento el agua inicialmente caliente y el agua inicialmente más fría alcanzaban la misma temperatura en su carrera hacia el punto de congelación, medí la temperatura aproximadamente cada cinco minutos durante los experimentos en que el par termoeléctrico estaba cerca del fondo. Me pareció particularmente interesante la comparación, minuto a minuto, entre la curva de enfriamiento del agua inicialmente a la temperatura de 81 grados Celsius (la que más tarda en alcanzar el punto de congelación) y la curva del agua inicialmente a una temperatura incluso superior: 95,5 grados. El agua que, en un principio, era la más caliente, alcanzaba la misma temperatura que la otra (la de 81 grados iniciales) a los siete u ocho minutos del comienzo de la refrigeración. Posteriormente, la temperatura del agua inicialmente más caliente se mantenía por debajo de la del agua inicialmente más fría hasta unos 15 o 20 minutos antes de la congelación. Poco tiempo antes de llegar a la congelación, la temperatura del agua inicialmente más caliente disminuía bruscamente alcanzando el punto de congelación, mientras que la del agua inicialmente más fría permanecía alrededor de un grado Celsius, durante unos minutos más.

Si la velocidad de enfriamiento fuera exponencial, es decir, si la temperatura del agua dependiera de una función exponencial del tiempo transcurrido, cuando se representaran los datos sobre papel semilogarítmico, darían una línea recta. Al llevar a cabo esta representación, comprobé que la curva de enfriamiento tenía más pendiente en los primeros y en los últimos minutos de cada experimento; en cambio, durante el resto del tiempo la temperatura variaba exponencialmente, pues la representación de los datos producía líneas rectas. Las muestras de agua que empezaban a 81 y 55 grados, respectivamente, presentaban dos regiones en su curva, puesto que, para cada una de ellas, tuve que utilizar dos líneas rectas con pendientes diferentes para unir los datos entre sí. A excepción de los primeros y de los últimos minutos, las curvas de ambas muestras eran paralelas, lo que parecía indicar que en ambos casos había mecanismos de enfriamiento similares. El agua cuya temperatura inicial era de 95,5 grados tenía una pendiente más fuerte en la primera mitad de la prueba, lo que sugería un mecanismo adicional de pérdida de calor que no tenía el agua cuya temperatura inicial era inferior. En otras palabras, por encima de una temperatura inicial de 81 grados Celsius, o bien las pérdidas por evaporación o bien la reciclación aumentan bruscamente de forma significativa y aumenta la velocidad del enfriamiento.

Las mediciones de temperatura tomadas cerca de la superficie del agua mostraron un pico más pronunciado en la curva del tiempo necesario para alcanzar la congelación que en el caso de las mediciones tomadas cerca del fondo. Pero, al parecer, la temperatura inicial del agua que más tardaba en helarse era la misma. Con



Pérdida por evaporación en el experimento hecho con 100 mililitros

una misma temperatura inicial de 70° C, hice dos pruebas para tomar las mediciones de temperatura tanto en la superficie como en el fondo y así ver cómo se diferenciaban dichas temperaturas durante el enfriamiento. Pocos minutos después del comienzo del ensayo, la superficie estaba unos pocos grados más fría. La diferencia aumentó a tres grados hacia la mitad de la prueba y después se redujo otra vez hasta unos pocos minutos antes de que la superficie empezara a congelarse. Sería interesante realizar más mediciones de este tipo, porque las corrientes de convección que se forman en el agua son originadas por la diferencia de temperatura entre la superficie del agua y el fondo. Quizá le guste al lector seguir trabajando sobre esta diferencia de temperatura para determinar la posible influencia de las corrientes de convección en la congelación. Pero le sugiero que en lugar de repetir los ensayos como yo hice, realice una sola prueba sumergiendo en el agua dos termopares con el fin de sacar mediciones de temperatura a diferentes profundidades, a la vez.

Para reducir el efecto de evaporación, por lo menos en lo que se refiere a eliminar la pérdida de masa debida a la evaporación, repetí el ensayo de 100 mililitros de agua en un recipiente de 600 mililitros, con el termopar cerca del fondo; pero esta vez puse sobre el borde del recipiente una bolsa de plástico y la sujeté al mismo con una banda elástica. La evaporación todavía restaba calor al agua, pero la masa evaporada se condensó en el plástico y cayó de nuevo al recipiente en forma de gotas. Para todas las temperaturas iniciales aumentó el tiempo necesario para la congelación. El resultado me indicó que la pérdida de masa durante la evaporación es importante.

Medir la pérdida de masa en ensayos normales deduciendo cuánta agua se necesita para rellenar el recipiente hasta su nivel original no es un proceso muy exacto. Sin embargo, el resultado general es que se pierde más agua cuando la temperatura inicial es más alta. En un estudio teórico de la velocidad de enfriamiento del agua, George S. Kell llegó a la conclusión de que enfriando el agua de 100 a cero grados Celsius la pérdida es de aproximadamente 16 por ciento de su masa, si la principal pérdida de calor es por evaporación. Mis resultados, según parece, coinciden con los cálculos de Kell, lo que nos confirma otra vez que la pérdida de calor y masa por evaporación es un factor importante en el paradójico efecto de congelación.

La recirculación interna del agua que se está enfriando puede ser examinada echando una pequeña cantidad de colorante en el recipiente de agua con una jeringa hipodérmica. El movimiento, a grandes rasgos, es de ascenso por el centro y de descenso por los bordes exteriores. El transporte de calor de esta recirculación está determinado probablemente, por la profundidad del agua, por el área de la sección transversal del recipiente y por la diferencia de temperatura entre la parte alta y la parte baja de la columna de agua.

No conozco ningún cálculo fácil que demuestre cómo influyen estos factores en el transporte de calor y, por tanto, no sé hasta qué punto el efecto de congelación se debe a la recirculación del agua.

¿Cómo influye la superficie en el tiempo necesario para alcanzar la congelación? ¿Existe una relación particularmente ventajosa entre el área de la superficie y el área lateral, que elevara el trazado de las curvas? Si se tuviera acceso a una cámara frigorífica, se podría entrar en ella para medir la pérdida de masa de los recipientes durante el enfriamiento llevando a cabo el experimento sobre el platillo de una balanza. La relación de pérdida de masa con la temperatura del agua en función del tiempo puede ser comprobada por medio de predicciones teóricas.

A Kell se debe una conjetura interesante: una de las razones por las cuales este efecto de congelación ha sido olvidado durante largo tiempo es que la sustitución de las cubetas de madera por las de metal disminuyen o eliminan el efecto. La conducción más rápida del calor a través de las paredes y del fondo del recipiente puede disminuir la importancia de la evaporación en el enfriamiento. Firth, sin embargo, está en contra de esta conclusión. El hecho se podría comprobar usando recipientes del mismo tamaño pero fabricados con distintos materiales, por ejemplo, metal y plástico. También se puede probar con dos recipientes de pirex, colocando una capa de aislamiento térmico alrededor de las paredes y por debajo de uno de ellos.

Los resultados acerca de la pérdida de masa pueden mejorarse cortando los recipientes justos por encima de la línea del agua. Cuando el recipiente se recorta de este modo, el agua evaporada no puede adherirse al mismo para volver a caer dentro de la misma.

Además de eliminar la pérdida de masa poniéndole una tapa al recipiente, también se puede trabajar vertiendo una fina capa de aceite en la superficie. Si el efecto de congelación se debe a la evaporación, lo lógico es que en este caso disminuya.

En mis experimentos, eliminé uno de los principales factores que influyen en la velocidad de enfriamiento del agua del grifo, esto es, el gas disuelto que acarrea dicha agua. Puesto que controlar la cantidad de gas disuelto en el agua sería muy difícil, el lector puede por lo menos intentar hacer el experimento siempre en las mismas condiciones si no deja de utilizar agua del grifo previamente hervida, ya que hay otra cuestión pendiente ¿retrasa realmente el enfriamiento del agua el gas que lleva disuelto?

También se podría estudiar qué ocurre con la velocidad de enfriamiento cuando se le añade sal al agua para rebajar el punto de congelación. Quizás el efecto sea más asombroso si en vez de agua se usara otro líquido (por ejemplo, alcohol) que se evapore más deprisa. El lector logrará congelar el alcohol a no ser que tenga acceso a medios extremos de enfriamiento, tales como nitrógeno líquido, pero a buen seguro el tiempo necesario para llegar a los cero grados manifieste la misma curva ascendente que se observa con el agua.

Libros

Escrituras, lenguas, los tres minutos iniciales del universo, calzadas romanas y quimeras

Philip Morrison

CORPUS OF MAYA HIEROGLYPHIC INSCRIPTIONS, VOLUME 1: INTRODUCTION TO THE CORPUS, por Ian Graham, Museo Peabody de Arqueología y Etnología, Universidad de Harvard. CORPUS OF MAYA HIEROGLYPHIC INSCRIPTIONS, VOLUME 2: PART 1, por Ian Graham y Eric von Euw, Museo Peabody de Arqueología y Etnología, Universidad de Harvard. THE BOOK OF A THOUSAND TONGUES, segunda edición dirigida por Eugene A. Nida. United Bible Societies, distribuido por Landmark Book Co., Nueva York. El empuje de la Ilustración y la fascinación del nuevo renacimiento de la cultura griega impulsaron a la Academia de Ciencias de Berlín a encargar a un destacado clasicista que formase y publicase una colección definitiva de todas las inscripciones remanentes de la Hélade. En 1882, August Böckh presentó el primer volumen de su gran *Corpus Inscriptionum Graecarum.* Hasta el día de hoy continúan apareciendo volúmenes de esa serie, que es un monumento de la epigrafía clásica. Pero el mundo supo hace poco tiempo algo sobre una cultura muy diferente, la cultura maya; toda la epigrafía maya es más reciente que los volúmenes clásicos de Böckh.

Hay dos fundadas razones para evocar el viejo *Corpus.* La creciente erudición y las nuevas herramientas –el ordenador– han dado una nota de optimismo acerca de los actuales empeños por descifrar los glifos mayas. Más que eso, los dispersos yacimientos mayas han constituido por más de una década la presa de ladrones organizados y aerotransportados, que se abren camino hasta los solitarios lugares y estropean y se llevan todo lo que sea atrayente y portátil. La obra de los mayas es atacada con sierras, taladros para granito, fuego e incluso explosivos. Muchas esculturas han sido destruidas para siempre por "saqueadores completamente carentes de pericia al tratar de adelgazar estelas para facilitar su transporte". Son hombres rudos. A veces han llegado a agredir físicamente al solitario guardián ¡Piratas! Son sólo el primer eslabón de una dorada cadena que une a marchantes "respetables", de Zürich a Los Angeles, y a través de ellos, a codiciosos coleccionistas e insensibles museos (¿antimuseos, quizá?).

Ian Graham se ha propuesto durante un decenio de estudio, promoción y duro trabajo, en el campo y en las fundaciones benéficas, la tarea de encontrar, describir y preservar sobre papel los textos mayas. Trabajo que continúa activamente. Estas primeras partes se publicaron en 1975 y a lo largo del año en curso aparecerán varias más. El plan contempla la introducción y cinco volúmenes, cada uno de ellos en tres partes, registrando en una zona geográfica tras otra todo texto maya conocido, excepto los escritos en alfarería o en papel: "jade, nácar, hueso, madera, estuco y muros pintados". Los relieves en piedra forman ciertamente el grueso de lo que tenemos; dondequiera se unan jeroglíficos y arte figurativo, se registrarán los diseños completos. La obra es magra, hasta seca: fotografías con cuidadosa indicación de su escala (algunas estereográficas), dibujos lineales meticulosos y mapas de emplazamientos, sin mayores comentarios.

El colega de Graham, von Euw, se unió al proyecto en 1971 para comenzar su "iniciación en los dobles misterios de los jeroglíficos y del arte de arrear mulas". Estos textos tropicales no se hallan ni en museos ni en la Acrópolis; en su mayoría yacen, difícilmente accesibles, en pantanos y selvas. Veamos el camino maderero de 18 kilómetros hasta Naranjo, en la frontera entre Guatemala y Belice. "Con cierto trabajo, este camino suele poder hacerse pasable para vehículos con tracción en las cuatro ruedas durante la estación seca." Las grandes páginas de estos libros baratos de 60 páginas presentan sus estelas en la escala 1:10. No serán muchos los lectores que quieran el *corpus* completo, pero no pocos disfrutarían poseyendo uno o dos yacimientos o lugares favoritos conforme vayan apareciendo los volúmenes correspondientes.

Pasamos de un erudito museo de textos en una lengua enigmática a una notable guía de campo del lenguaje en acción. En este caso, con una modesta pero experta dotación de comentarios históricos y lingüísticos, se nos presenta la constatación de Babel. *The Book of a Thousand Tongues* pone ante el lector unas 1400 muestras, en un lenguaje tras otro, principalmente del mismo texto, tomado del Evangelio según San Marcos. Una cita identificará ese texto, legible para la mayoría de los angloparlantes, tomada de la lengua disciplinada y expresiva del golfo de Guinea, y denominada –en el libro– inglés pidgin del Camerún, pero conocida también como wescos (West Coast = Costa Occidental) y Krio (la núm. 1051 de las 1399 lenguas anotadas): "1-Di fos tok fo digud nyus fo Jesus Christ God yi Pikin. 2.I bi sem as i di tok fo di buk fo Isaiah". El volumen informa, al año 1970, sobre el estado de esa extraordinaria empresa evangélica, que continúa la labor iniciada en la comunidad judía de Alejandría, cuyos miembros hicieron poner, durante el siglo III a. C., su Testamento hebreo en buen griego helenístico. En el siglo IX d. C., san Cirilo ideó su alfabeto para verter el texto de la Biblia al ruso, tal como lo hicieron antes que él Mesrop para los armenios y Wulfila para los visigodos. Podemos ver aquí los versículos de Marcos en cirílico (versiones antigua y nueva), en armenio (una versión antigua y dos modernas, una para el dialecto de la República Armenia soviética y otra para el de la diáspora) e incluso en el antiguo alfabeto de Wulfila (muy desconectado del tipo de letra antigua que llamamos gótica).

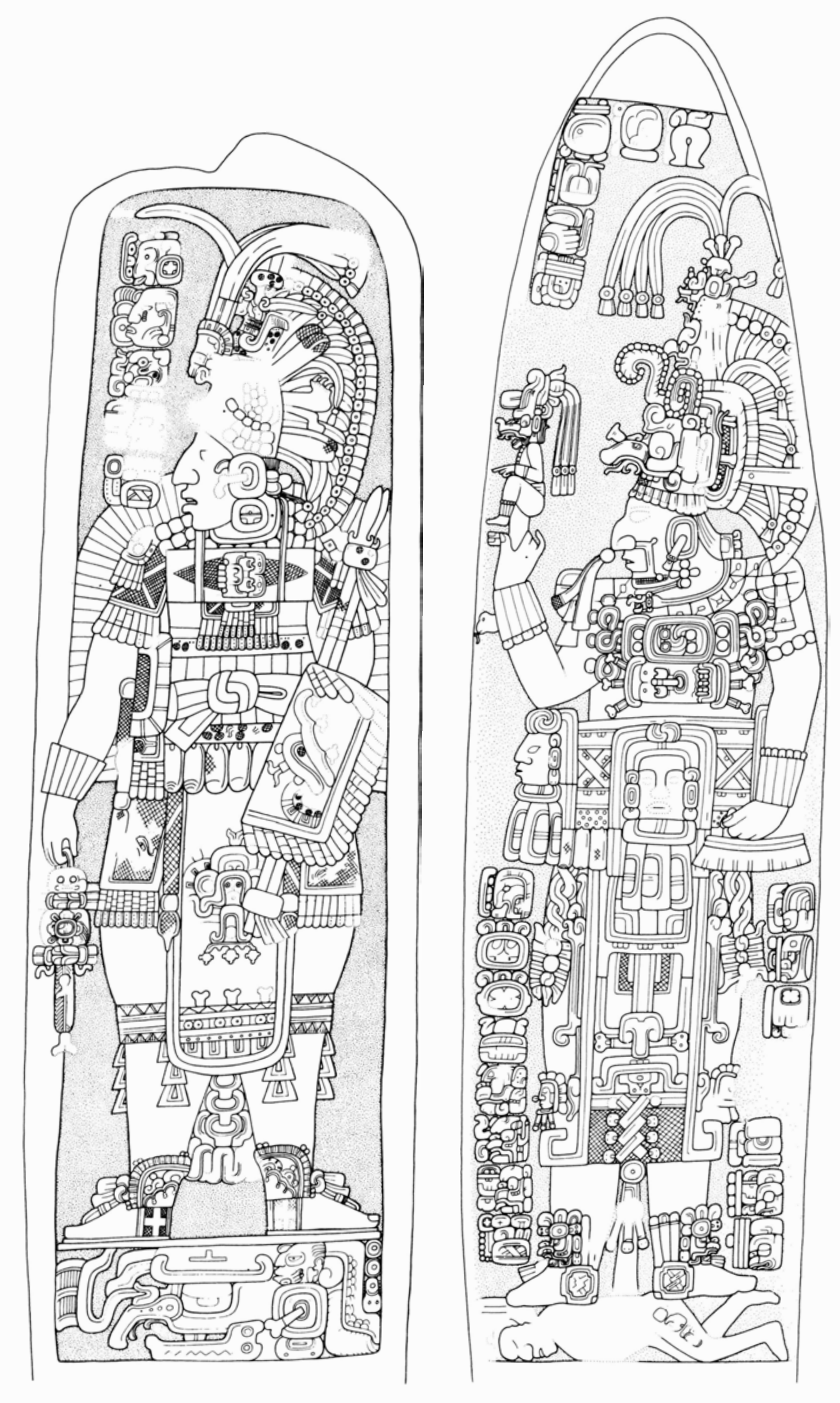
¿Chino? Ciertamente, y en un par de docenas de versiones: en antiguos y nuevos caracteres de letra; romanizaciones; fonética en chino, construida, al igual que el silabario japonés, a partir de fragmentos de caracteres; dialectos para muchas provincias, y una cantidad de lenguajes de minorías no-Han de China. El texto en dakota, quechua, kiswahili, gitano, manchú y en el pitjantjatjara de los cazadores del desierto de Australia occidental podría servir para señalar los continentes. He allí los grandes lenguajes, como el árabe, en una docena de variedades, y también los pequeños. Véase por ejemplo el de los sirionó, antaño los "nómadas de arco largo" de las cabeceras del Amazonas, unos 5000 en total, asentados ahora en plantaciones, o el suki, lenguaje de un pueblo papú, "notorios cazadores de cabezas hasta los años cuarenta". Hay ejemplos de fenómenos lingüísticos generales: las *lingua franca,* tales como el police motu (Port

Moresby) y el kituba (Bajo Congo); las antiguas lenguas que han llegado a ser sagradas, tales como el latín y el sánscrito (en el censo de 1961 en la India "unas 2500 personas mencionaron el sánscrito como su lengua materna"); los lenguajes inventados, como el esperanto y el volapük (contrapartida alemana de aquél).

En las montañas del sudoeste de China y por las fronteras de Laos, Tailandia y Vietnam hay esparcidos un par de millones de personas conocidas hace ya largo tiempo (por los chinos) como los miao negros, blancos, azules, rojos o floridos. La distinción no procede del color de su piel, sino de sus trajes característicos; las mujeres miao son mundialmente famosas por la calidad artística de sus telas. Sus textos están en escritura pollard, un silabario confeccionado por la misión allí existente durante el reinado de Eduardo VIII (1901-1910). En Canadá, hace más de un siglo, otro émulo de san Cirilo dio pasos heroicos hacia un evangelio impreso en lengua cree. La obra de James Evans es una leyenda: no tenía papel, tinta, tipos ni imprenta. Talló en madera los símbolos que inventó, estampando moldes de arcilla con sus troqueles de madera. En esos moldes derramó plomo derretido, metal obtenido del revestimiento de cajas de té; su prensa estaba hecha de piel, su tinta era aceite de pescado y su papel, corteza de abedul. El libro estaba cosido con tiras de cuero y encuadernado con piel de venado.

Este conjunto de ejemplos de lenguajes es una colección de trabajo notable: cada texto está dirigido a lectores vivientes, de modo que la difícil distinción entre lenguajes y dialectos se enfrenta más o menos operacionalmente. Menos de una décima parte de las listas son de dudosa individualidad: la mayoría de ellas son los textos escritos en más de 60 dialectos de las principales lenguas europeas que se prepararon allá por 1860 para el príncipe Louis Lucien Bonaparte, sobrino de Napoleón. Son antiguallas de escasa utilidad.

La devota labor de los misioneros ha puesto, por estas fechas, parte de la Biblia en el idioma nativo de alrededor del 97 por ciento de los pueblos de la tierra. La tarea no es sólo de traducción sino también de impresión. A los misioneros, el trabajo no les parece completo. Quedan todavía unos 1000 lenguajes por andar e "incluso ahora hay traductores trabajando en más de 500 de esos lenguajes". La rentabilidad decreciente en cuanto a nuevos lectores es obviamente visible. En 1965, a otros 25 lenguajes se les dio su parte de la Biblia, pero ninguno de ellos tenía más usuarios que el kizaramo, lengua bantú de Tanzania, con un millón de personas, y la mayoría eran idiomas locales menores, como los de los altiplanos políglotas de Nueva Guinea o de las remotas montañas de Oaxaca.



Dos estelas mayas, del **Corpus of Maya Inscriptions**

Pueden trazarse cuatro períodos. Los primeros cristianos tradujeron a los idiomas dominantes de las fronteras militares del mundo antiguo: copto, siríaco, etíope, geórgico y semejantes. La Reforma vio el ascenso de los lenguajes nacionales de Europa. Luego vinieron los misioneros clásicos, que se diseminaron con el imperio por todo el mundo, llegando a los indios americanos, a Persia y Bengala, a China y Japón. Ahora hay una cuarta oleada, que evidentemente se apoya más en nacionales educados que

en extranjeros llenos de celo evangélico. La primera traducción bíblica a un idioma africano principal "llevada a cabo por cristianos nativos" fue la Biblia yoruba, terminada bajo la dirección de Samuel Ajayu Crowther, obispo de Nigeria en 1850-59. Crowther había sido vendido como esclavo de muchacho, pero el negrero fue interceptado por patrullas navales inglesas y Crowther fue devuelto a Freetown, Sierra Leona, en donde concibió el "sueño de volver y evangelizar a su propio pueblo".

El asombroso número de 70 diferentes escrituras aparece en este libro, más otros 75 lenguajes que usan la escritura romana agregándole un número considerable de letras especiales. Hay bastante ironía en esta crónica del fervor evangélico. He ahí un texto facsímil, con las letras bastantes rotas, impreso en 1685 en la imprenta de Samuel Green, que había sido presidente del Harvard College. Lleva el pasaje de San Marcos en traducción de John Eliot de Roxbury, quien trabajó 15 años para aprender el idioma de los massachusetts. "Yo había tenido el ardiente deseo, si pluguiese a Dios, de que nuestro lenguaje indio pudiese ser santificado al traducir a él las Sagradas Escrituras." Eliot perseveró, y la primera edición de su obra en 1663 se convirtió en la primera Biblia impresa en América del Norte. Los massachusetts recibieron por fin la buena nueva, pero para entonces la viruela, el hacha y las armas de fuego les habían llevado noticias de índole muy distinta. Son pocos quienes pueden leer hoy la Palabra en lengua massachusetts.

THE FIRST THREE MINUTES: A MODERN VIEW OF THE ORIGIN OF THE UNIVERSE, por Steven Weinberg. Basic Books Inc. Publishers. THE RED LIMIT: THE SEARCH FOR THE EDGE OF THE UNIVERSE, por Timothy Ferris. William Morris and Company, Inc. La vaca Audhumla dio forma a los primeros seres humanos lamiendo el hielo primordial. Es la cosmología nórdica del *Edda Menor* la que informa de los primeros párrafos del libro compacto y lúcido del profesor Weinberg, recia destilación de la "gran explosión" de la cosmología actual. En ocho breves capítulos, dirigidos a un lector que sea "un hábil abogado viejo que no habla mi idioma, pero... espera... escuchar algunos argumentos convincentes". Weinberg bosqueja rápidamente la expansión del universo de galaxias, da cuenta del descubrimiento del trasfondo uniforme de microondas, y sobre eso construye, con unos pocos gráficos y números, pero sin ecuaciones, un relato detallado de la ardiente y tremenda explosión. Pasa ante nuestros ojos un filme en palabras, seis cuadros no espaciados parejamente en el tiempo sino más bien en temperatura, cada uno de los cuales muestra un universo tres veces más frío que el precedente.

El filme comienza –lo que sería algo después del inicio de la materia y del tiempo –en una caliente sopa universal en expansión hecha de electrones y positrones, fotones y neutrinos, y el rociado de nucleones que componen la materia del mundo presente. En el segundo cuadro, una décima de segundo más tarde, las cosas continúan enfriándose, y poco ha cambiado. En el tercer cuadro, los neutrinos vuelan, liberados de la materia, la cual se ha adelgazado demasiado como para atraparlos. En el cuadro siguiente, cuando el relato no lleva ni 15 segundos, los electrones y positrones se están recombinando en radiación, y se están formando deuterones a partir de choques entre protones y neutrones. Aún hay demasiado calor como para que persistan los deuterones. En el quinto cuadro quedan sólo fotones y los neutrinos libres; los neutrones están empezando a desintegrarse y, poco después de este cuadro, a una temperatura unas decenas de veces mayor que la del centro de nuestro sol actual, los deuterones pueden mantenerse unidos. Los deuterones que chocan con protones o neutrones, o entre sí, se unen formando núcleos de helio, en una o dos etapas. Los neutrones restantes se cocinan así, rápidamente, para constituir helio, pero no puede formarse un núcleo más pesado; el helio no puede recoger neutrones o protones para formar un núcleo estable, y no es posible que dos núcleos de helio se unan entre sí. La síntesis de materia se detiene en el helio hasta que, largo tiempo después, en el centro denso y duradero de ciertas estrellas a veces pueden encontrarse unos con otros tres núcleos de helio. Nace el carbono y, mucho después, la vida.

En el modelo académico –los números son, por supuesto, inciertos– han pasado tres minutos cuarenta y seis segundos desde el primer cuadro. El por qué habríamos de creer este cuento, más extraño que el de Audhumla, y en dónde podría estar equivocado (hay ciertos problemas referentes a la isotropía que aún nos intrigan), es el tema del resto del libro. El catedrático Weinberg es uno de los teóricos más productivos de la actual física norteamericana; se las ha arreglado para transmitir ese gusto curiosamente sabroso del cálculo específico y coherente, que deleita al teórico. La teoría quizá sea errónea, pero es lo bastante exquisita y cuantiosa como para ser definida. Hay mayores dudas acerca de los primeros momentos; la primera centésima de segundo se examina a la luz de especulaciones esperanzadas nacidas de la física de partículas, física en curso de aparición. La profunda simetría de las partículas, hoy quebrantada y obscura, puede haber sido íntegra entonces. En esos tiempos los "fenómenos físicos exhibían directamente la simplicidad esencial de la naturaleza. Pero nadie había para verla". Qué está por venir aún –una dilución interminable, o un regreso al horno ardiente –no está claro todavía.

A este librito se ha agregado un glosario, una cuidadosa bibliografía a diversos niveles de tecnicismo y una docena de páginas de suplemento matemático, de manera que quienes quieran las fórmulas algebraicamente sencillas, pero profundas, en las cuales se apoyan los argumentos, puedan ver la base de la cosmología.

The Red Limit no es obra de un teórico, sino de un periodista científico caviloso y penetrante. Se inclina hacia el lado personal de la cosmología moderna tanto como Weinberg hacia lo lógico y lo cuantitativo, aunque ambos autores se unen para narrar los accidentes y personalidades que rodean el descubrimiento clave del trasfondo universal de radiación de microondas, un relato tan enredado como el que más de los que pueda hallarse en la ciencia sobre oportunidades perdidas, predicciones pasadas por alto, experimentación laboriosa y, al final, conclusiones grandiosas e inconmovibles. Ferris escribe con auténtica gracia, y ha entrevistado a casi una veintena de los hombres de telescopio, amplificador y ecuación, que representan la gama completa de la cosmología contemporánea en los Estados Unidos. Constituye un placer y una revelación leer qué les pasó a Maarten Schmidt y Jesse Greenstein un día de febrero de 1963; Schmidt caminaba, corredor abajo, a la oficina de Greenstein, en el Instituto de Tecnología de California, para explicar que los enigmáticos espectros de quasar que los habían traído de cabeza a ambos durante un año, mostraban simplemente desplazamientos nunca vistos hacia el rojo. La solución de Schmidt incitó a Greenstein a una inmediata concordancia: en algún lugar de su mente ya había llegado a la conclusión correcta sin saberlo conscientemente, y en el acto enunció el desplazamiento análogo, correcto hasta dos decimales, para un ejemplo en particular, antes de que verificaran esa placa en los archivos. He allí a Halton Arp, escéptico cosmólogo, que elige los objetos para el telescopio de 508 centímetros "con ojo de artista". Así como Allan Sandage –"un hombre

de absoluta integridad", ferviente continuador del gran programa de Hubble-busca una constancia entre las galaxias, Arp persigue siempre la novedad.

En el libro de Ferris hay muy pocos números y ningún gráfico; expone los fundamentos de la visión cosmológica moderna a lo ancho de un amplio intervalo histórico, desde Friedmann y Hubble hasta Wheeler y Hoyle. La calidad íntima de su reportaje y el reflexivo encuadramiento en que el autor dispone los temas, señalan como excepcional a este libro, especialmente para lectores distanciados de las ciencias físicas. Estos dos libros tan diferentes, de autores de historial y personalidad tan distintos, terminan en una nota concordante, con un eco de Pascal. Para Weinberg, el esfuerzo por comprender el universo presta a nuestra vida humana "algo de la gracia de la tragedia". Para Ferris, un poco menos melancólico, la inteligibilidad de la naturaleza es una afirmación de fe. Cita a A. S. Eddington: "No hay nada de erróneo en ello. Después de todo, somos parte del universo... Esa fe es también parte del universo".

LES VOIES ROMAINES, por Raymond Chevallier, Colin, París. "Todos los caminos conducen a Roma", dicen. No se sabe exactamente de quién procede el dicho; ¿Cicerón, tal vez? Este estudio erudito y ameno de un profesor de latín de Tours constituye una lectura admirable por su vivacidad, nada diferente del viaje a lo largo de la Vía Appia hasta Brindisi, descrito por Horacio, lleno de alegres paradas, riñas ocasionales, pequeñas decepciones y mucho por aprender.

Había 87.000 kilómetros de carreteras principales: "El poder de Roma ha dado unidad al mundo", escribió con orgullo Plinio el Viejo. Las carreteras eran un instrumento primordial de esa unidad, política y militar en primera instancia. Chevallier dice que la zona romana fortificada que cubre el sur de Argelia, de 60 a 80 kilómetros de espesor por 300 de anchura, con caminos de enlace y pequeños fuertes abandonados por miles, fue un proyecto "más grandioso que la Muralla China o la Línea Maginot". Su estudio y análisis debe mucho al experto coronel Baradez, que la investigó, tal como el padre Poidebard (asimismo oficial de la reserva), pionero del uso de la aerofotografía en arqueología, investigó la red análoga existente en el desierto de Siria.

Estos estudios de los *limites* dicen mucho acerca del sistema romano. Los caminos romanos suelen ser rectos, construidos no sobre la divisoria de las

colinas y elevaciones, sino inmediatamente bajo esa línea, en la "cresta militar" en donde las tropas pueden marchar sin ser vistas pero pueden todavía ver lo suficiente por encima de dicha cresta. Abajo, por la carretera principal, se apresuraban o haraganeaban los viajeros o el comercio del imperio, a caballo, en cómodas literas llevadas por hombres o mulas, en calesines o coches, sobre pesados carros arrastrados por numerosas yuntas, o a pie. Las legiones andaban 10.000 pasos por día, 15 kilómetros, en marcha regular y doble cuando estaban en *magnum iter*. (A los soldados se les daba un día de descanso por cada cuatro.) Con bagaje pesado –la "artillería"– la tropa iba con mayor lentitud. El correo imperial llevaba a los funcionarios y sus mensajes en sillas de posta tiradas por caballos, las cambiaba en relevos preparados, a la velocidad de unos 75 kilómetros por día. Había tarjetas diplomáticas de crédito, de bronce, y también falsificaciones.

El meollo de este libro es la naturaleza de las pruebas. Las hay de toda especie. Están los escritos de poetas y diaristas, la legislación, las historias (que a menudo no sirven gran cosa a causa de su énfasis político). Luego están las inscripciones de las piedras miliares uniformes. Contamos con unas 4000 piezas miliares inscritas en latín, de las cuales más de la mitad están en Africa. Pocas se mantienen en su sitio; ha sido preciso buscarlas en las aldeas, en donde a lo largo de los siglos han servido como finos rodillos o muelas para afilar hachas o, un poco ahuecadas, como pilas para agua bendita. No es tarea fácil hacer concordar las piedras marcadas con los diversos itinerarios y mapas que poseemos procedentes de Roma, unos en copias medievales, otros inscritos en cálices de plata o en copas esmaltadas. Nombres toponímicos y leyendas de peregrinos, algunas señales camineras (de autenticidad incierta) y vidas de santos conservan sobre papel las viejas calzadas.

El terreno es por cierto la fuente más segura, pero no es fácil. Los caminos y puentes de ciudades dan indicios, las puertas de ciudades son encrucijadas, las lindes de los campos preservan alineaciones largo tiempo después de haber desaparecido la carretera misma. Los campos de Galia declaran los viejos caminos, incluso en mapas y en recientes escrituras y límites de propiedades. Naturalmente, la aerofotografía es una poderosa herramienta. Es evidente que las calzadas seguían en general una línea recta, pero no ciegamente; muestran una práctica racional de la ingeniería, "respetando la disposición del terreno", de manera que a veces la línea se curva. La construcción no consistía en la sencilla pared enterrada, tan querida para los libros de texto. Las carreteras siempre se cimentaban y drenaban cuidadosamente, se construían de muchas capas y no siempre se pavimentaban. Los constructores prestaban mucha atención a las condiciones locales, en especial a las fuentes de materiales y a las tierras pantanosas. Recurrían con frecuencia a terraplenes. En donde hay corrientes de agua por cruzar se encuentran vados, estaciones para las balsas de los profesionales encargados del trasbordo y, en las carreteras más grandes, puentes.

Las fotografías aéreas del paisaje europeo frecuentemente exhiben las carreteras romanas. Sin embargo, hay trampas para el incauto. Una vía férrea de trocha angosta abandonada puede parecer un camino romano; tuvo que solucionar los mismos problemas que aquél. La vía busca un curso rectilíneo, hace buen uso de taludes, evita los pequeños asentamientos y corre junto a una carretera principal de nuestros días. Sólo la amplia curva que debe seguir en las cuestas la traiciona ante el fotointerpretador: "No hay nada más parecido a un camino romano en su trazado, color, anchuras y aspecto general".

Este no es tanto un libro de referencia cuanto una exposición copiosa y discursiva hecha por un hombre enamorado del problema; su bibliografía cuidadosamente clasificada debería servir de ayuda a quien quiera más. El lector común disfrutará del viaje. Las ilustraciones son frescas y convincentes, aunque los mapas camineros del imperio son pequeños.

MAMMALIAN CHIMAERAS, por Anne McLaren, F.R.S. Cambridge University Press. REGENERATION, por Priscilla Mattson. The Bobbs-Merril Company, Inc. Un padre y una madre aportan por igual a las numerosas propiedades innatas de cada uno de nosotros. Nos desarrollamos a partir de un solo huevo, esa célula recién dotada de un paquete de cromosomas paternos. Crecemos mediante un desfile suntuoso, dinámico e intrincado de fisiones y cuasi fusiones a través de 40 etapas de división celular hasta llegar a ser un organismo adulto desarrollado; "cada mamífero es una labor de retazos hecha de diferentes clases de células". Si se corta la piel o se quiebra un hueso, el proceso vuelve a empezar localmente, y crecen nueva piel o nuevo hueso hasta casi alcanzar su estado original. Si un tritón pequeño pierde incluso una pata completa, al cabo de dos meses otra pata está de vuelta en su sitio.

La intervención experimental en este proceso data de antiguo; se remonta al viejo Lazzaro Spallanzani y sigue siendo vigorosa como especialidad experimental en todo el mundo, especialmente en la URSS. En la regeneración, prácticamente se vuelve por segunda vez a retomar la huella o pista del desarrollo. Y, lo que es aún más notable, la pista completa puede ser iniciada de modo inesperado. La quimera es un animal compuesto, de tamaño y forma normales, hecho a retazos de dos (o más) poblaciones distintas de células, en el que cada conjunto suele tener sus dos propios progenitores. Experimentalmente, el fenómeno tiene unos 15 años de edad; ha atraído a unas docenas de investigadores especializados, permanentemente fascinados por las "preguntas que nadie soñó jamás que existiesen en los días en que un individuo tenía sólo dos progenitores".

En el gran museo de Florencia el visitante puede ver una hermosa quimera con signos de insólita regeneración. Es un león con cola de serpiente, cuello y cabeza de cabra que nacen de su lomo, obra en bronce de un anónimo artista etrusco (reparada con mano experta por Cellini). En *Mammalian Chimaeras*, monografía de ritmo rápido, pero sumamente técnica (y con una amplia bibliografía), escrita por un inglés destacado en el tema, el lector puede ver numerosas quimeras de carne y hueso, tanto ratas como hombres. En *Regeneration*, librito bien informado pero sencillo y popular obra de un joven investigador de Western Case Reserve, vemos una extraña extremidad supernumeraria de tritón, entre una variedad de estructuras regeneradas, desde medios platelmintos hasta cornamentas de ciervo. El tejido en crecimiento resulta ser casi tan plástico como el bronce, una vez que se aprende el oficio. Los dos libros, en sus diferentes estilos, son apasionantes, y ambos dejan muy en claro que acerca de cómo funciona todo eso no sabemos mucho más que lo que sabía el hábil etrusco sobre los diagramas de fases del cobre y sus aleaciones.

Si se quiere inducir en una salamandra un miembro no especificado en el plano, el modo de hacerlo es claro. Córtese una región de músculo pectoral, ocasionando una lesión histológica como la que queda en el muñón de una pata cortada. Llévese hasta esa zona algunos nervios importantes de las extremidades, dejando libres sus extremos por corte, y póngaselos en contacto con los músculos desgarrados. En "cuarenta días una pequeña pata extra, con su esqueleto y musculatura bien conformados, se había establecido perfectamente". Incluso a algunos mamíferos, zarigüeyas recién nacidas, se les puede inducir a regenerar patas traseras

mediante un injerto de tejido nervioso en el lugar de la amputación. Alguna substancia liberada por las células nerviosas (sabemos algo acerca de qué no es) es indispensable para la regeneración. Para saltar a un linaje muy diferente, una sola célula del pelo radical de una zanahoria, de ningún modo embrionaria, fue criada hace una docena de años hasta que llegó a ser una planta de zanahoria floreciente y plenamente desarrollada, habiendo sido desdiferenciada dicha célula al cultivarla en un medio que contenía leche de coco, substancia relacionada con el embrión. Después se ha descubierto que el espárrago y el tabaco exhiben la misma capacidad, y puede emplearse un medio sintético completamente conocido. Está muy lejos de haberse probado la extrapolación obvia a cualquier célula aislada de cualquier organismo (cada una de ellas lleva la cinta completa del ADN del organismo en cuestión).

El ratón hijo de cuatro progenitores es un *tour de force* más reciente de los biólogos. La técnica normal se aplica en los ratones mediante la extracción de los embriones más tempranos, cada uno de los cuales es una masa microscópica que consta quizá de ocho células. Dos de dichas masas de células (el asunto funciona incluso hasta la etapa de 32 células) se tratan con un enzima que quita una delgada membrana protectora. Al empujarlos suavemente uno contra otro, los dos embriones se confunden en una sola masa. Después de uno o dos días de cultivo *in vitro,* la masa en crecimiento se traslada con delicadeza al útero de una ratona madrastra. A su debido tiempo nace un ratoncito; en un gran número de casos sobrevive y madura hasta convertirse en un ratón adulto de tamaño y forma normales. Pero ese ratón tiene una curiosa duplicidad en su constitución: en su piel (que a veces es un tanto a rayas) hay dos poblaciones de células genéticamente distintas. Se han empleado también ratas, conejos y ovejas con diverso éxito. Ningún ratón-rata u otra mezcolanza de especies ha sobrevivido más allá del nacimiento. En seres humanos se conoce la existencia de quimeras espontáneas; se ha demostrado un doble legado genético de al menos uno de los progenitores. Hay una gran cantidad de posibles mecanismos, siendo el más verosímil la fertilización, por un segundo espermatozoide, del cuerpo polar, de menos tamaño y generalmente eliminado, que se forma junto con el óvulo, y la conglomeración ulterior de ambas líneas para formar un embrión.

La quimera constituye sin duda una aguzada herramienta para la investigación. Esa labor está adviniendo con rapidez; la tarea consiste en usar las distintas poblaciones celulares entremezcladas para revelar los mecanismos del desarrollo. Se precisan marcadores de células; pueden ser un pigmento visible, cromosomas determinadores del sexo, antígenos, genes inductores de tumores o configuraciones de genes que de ordinario no se expresan, pero que pueden ser "rescatadas" por la actividad de las parejas respectivas de los genes no expresados. Las quimeras de ratón que contienen diferentes genes para el color del pelaje muestran en éste franjas transversales claras y oscuras, cada una de las cuales ocupa nítidamente una u otra mitad del lomo del animal, a derecha o a izquierda de la columna vertebral. La sugerencia inicial, según la cual se ve allí las decisiones primitivas adoptadas al convertirse una u otra de las células aisladas en la determinante de color al formar clones para toda una parte segmentada del animal, parece ser demasiado simplista. La división lateral sugiere que los clones de célula de pelaje estaban fijados antes de que se hubiesen fundido los pliegues neurales del embrión: a esas alturas ninguna célula podía atravesar la línea central. El número y la distribución de las franjas parecen demasiado complejos como para apoyar el modelo más sencillo.

Los espermatozoides de quimeras se parecen a los de una u otra cepa componente, sin formas intermedias. En cambio, las retinas de quimeras de pigmento muestran un entresijo de células pigmentadas y despigmentadas; el tamaño y la distribución de los "parches" retinales parecen exigir la adherencia de células, muchísima mezcla aleatoria y cierto número de divisiones celulares durante el desarrollo. Los períodos alternos de mescolanza celular y de proliferación celular son bastante generales; "sucesos muestrales" adjudican luego paso a paso los grupos de células a los nuevos órganos y tejidos. Si la mezcla precede al crecimiento clónico, todos los "parches" serán puros, con un marcador de pigmento claro u obscuro. En caso contrario, se prevén parches tanto puros como abigarrados. Se tiene la esperanza de hallar muchísimos marcadores para etapas intermedias y para estructuras especiales. En los mamíferos, las células no son totalmente independientes; se muestran prolíferas y migran; a modo de contraste, en la mosca de la fruta las células inician clones para ir construyendo tejidos y órganos, y la progenie de las células tiende a permanecer en su sitio. Una mosca es un grupo de aldeas montañesas estables y asentadas desde antaño; un mamífero es una población metropolitana móvil, interactuante y cambiante.

Bibliografía

Los lectores interesados en una mayor profundización de los temas expuestos pueden consultar los trabajos siguientes:

MICROELECTRONICA

MICROCOMPUTERS / MICROPROCESSORS: HARDWARE, SOFTWARE, AND APPLICATIONS, John L. Hilburn y Paul M. Julich. Prentice-Hall, Inc., 1976.

SPECIAL ISSUE ON MICROPROCESSOR TECHNOLOGY AND APPLICATIONS. *Proceedings of the IEEE,* vol. 64, n.º 6; junio, 1976.

ELEMENTOS DEL CIRCUITO MICROELECTRONICO

PHYSICS AND TECHNOLOGY OF SEMICONDUCTOR DEVICES. A. S. Grove. John Wiley & Sons, Inc., 1967.

THE THEORY AND PRACTICE OF MICROELECTRONICS. Sorab K. Ghandhi. John Wiley & Sons, Inc., 1968.

MICROPOWER CIRCUITS. James D. Meindl. John Wiley & Sons, Inc., 1969.

BASIC INTEGRATED CIRCUIT ENGINEERING. Douglas Hamilton y William Howard. McGraw-Hill Book Company, 1975.

INTEGRACION A GRAN ESCALA EN LOS CIRCUITOS MICROELECTRONICOS

DESIGN OF DIGITAL COMPUTERS. Hans W. Gshwind y Edward J. McCluskey. Springer-Verlag, 1975.

MICROCOMPUTERS. André G. Vacroux en *Scientific American,* vol. 232, n.º 5, págs. 32-40; mayo, 1975.

PHYSICS OF COMPUTER MEMORY DEVICES. S. Middelhoek y P. Dekker. Academic Press, Inc., 1976.

THE SMALL ELECTRONIC CALCULATOR. Eugene W. McWhorter en *Scientific American,* vol. 234, n.º 3, págs. 88-98; marzo, 1976.

LA FABRICACION DE MICROCIRCUITOS ELECTRONICOS

MOS INTEGRATED CIRCUITS; THEORY, FABRICATION, DESIGN AND SYSTEMS APPLICATIONS OF MOS LSI. Engineering staff of American Micro-systems. Inc., dirigido por William M. Penny y Lillian Lau. Van Nostrand Reinhold Company, 1972.

MOS/LSI DESIGN AND APPLICATION. William N. Carr y Jack P. Mize, dirigido por Robert E. Sawyer y John R. Miller. McGraw-Hill Book Company, 1972.

METAL-OXIDE-SEMICONDUCTOR TECHNOLOGY. William C. Hittinger en *Scientific American,* vol. 229, n.º 2, págs. 48-57; agosto, 1973.

MEMORIAS MICROELECTRONICAS

RANDOM-ACCESS MEMORIES. Rein Turn en *Computers in the 1980s.* Columbia University Press, 1974.

DIGITAL MEMORIES. Carlo H. Séquin y Michael F. Tompsett en *Charge Transfer Devices.* Academic Press. Inc., 1975.

MAGNETIC BUBBLE TECHNOLOGY: INTEGRATED-CIRCUIT MAGNETICS FOR DIGITAL STORAGE AND PROCESSING. Dirigido por Hsu Chang. IEEE Press, 1975.

SPECIAL ISSUE ON LARGE CAPACITY DIGITAL STORAGE SYSTEMS. *Proceedings of the IEEE,* vol. 63, n.º 8; agosto, 1975.

NEW MEMORY TECHNOLOGIES. Jan A. Rajchman en *Science,* vol. 195, n.º 4283, págs. 1223-1229; 18 marzo, 1977.

MICROPROCESADORES

SPECIAL ISSUE ON MICROPROCESSOR TECHNOLOGY AND APPLICATIONS. *Proceedings of the IEEE,* vol. 64, n.º 6; junio, 1976.

SPECIAL ISSUE ON SMALL SCALE COMPUTING. *Computer,* vol. 10, n.º 3; marzo, 1977.

TRENDS IN COMPUTERS AND COMPUTING: THE INFORMATION UTILITY. Stuart E. Madnick en *Science,* vol. 195, n.º 4283, págs. 1191-1199; 18 marzo, 1977.

MICROCOMPUTADORES

MICROCOMPUTER-BASED DESIGN. J. B. Peatman. McGraw-Hill. Inc., 1976.

MICROPROCESSOR SYSTEM DESIGN. E. E. Klingman. Prentice-Hall, Inc., 1977.

SOME ASPECTS OF THE IMPLEMENTATION ON A MICROCOMPUTER OF SEQUENTIAL SYSTEMS A. Alabau y J. Figueras, en *Proceedings of Minicomputers and Microcomputer Conference,* Zurich; 2-5 junio, 1975.

ALGORITMOS. D. Knuth en *Investigación y Ciencia,* n.º 9, págs. 42-53; junio, 1977.

MICROELECTRONICA Y PROCESO DE DATOS

MINICOMPUTER SYSTEMS: STRUCTURE, IMPLEMENTATION AND APPLICATION. Cay Weitzman. Prentice-Hall, Inc., 1974.

DATA PROCESSING IN 1980-1985: A STUDY OF POTENTIAL LIMITATIONS TO PROGRESS. T. A. Dolotta. M. I. Bernstein, R. S. Dickson, Jr., N. A. France, B. A. Rosenblatt, D. M. Smith y T. B. Steel, Jr. John Wiley & Sons, Inc., 1976.

TRENDS IN COMPUTER HARDWARE TECHNOLOGY. David A. Hodges en *Computer Design,* vol. 15, n.º 2, págs. 77-85; febrero, 1976.

MICROELECTRONICA E INSTRUMENTACION Y CONTROL

ELECTRONIC MEASUREMENTS AND INSTRUMENTATION. Dirigido por Bernard M. Oliver y John M. Cage. McGraw-Hill Book Company, 1971.

MICROPROCESSORS: DESIGN AND APPLICATIONS IN DIGITAL INSTRUMENTATION AND CONTROL. Walter Banks y Jayanti C. Majithia en *IEEE Transactions on Instrumentation and Measurement,* vol. IM-25, n.º 3, págs. 245-249; septiembre, 1976.

DIGITAL SYSTEMS SPAWN NEW TASKS IN MEASUREMENT. A. Santoni en *Electronics,* vol. 49, n.º 22, págs. 100-106; 28 octubre, 1976.

LSI CHIPS TAKING OVER MORE HOUSEHOLD CHORES. G. M. Walker en *Electronics,* vol. 49, n.º 22, págs. 128-134; 28 octubre, 1976.

MICROELECTRONICA Y COMUNICACIONES

DIGITAL CODING OF SPEECH WAVEFORMS: PCM, DPCM, AND DM QUANTIZERS. Nuggehally S. Jayant en *Proceedings of the IEEE,* vol. 62, n.º 5, págs. 611-632; mayo, 1974.

OPTICAL TRANSMISSION OF VOICE AND DATA. Ira Jacobs y Stewart E. Miller en *IEEE Spectrum,* vol. 14, n.º 2, págs. 32-41; febrero, 1977.

SINGLE-SLICE SUPERHET. William Peil y Robert J. McFadyen en *IEEE Spectrum,* vol. 14, n.º 3, págs. 54-57; marzo, 1977.

MICROELECTRONICA E INFORMATICA

FUNDAMENTAL LIMITATIONS IN MICROELECTRONICS, I: MOS TECHNOLOGY. B. Hoeneisen y Carver A. Mead en *Solid-State Electronics.* vol. 15, n.º 7, págs. 819-829; julio, 1972.

LIMITATIONS IN MICROELECTRONICS, II: BIPOLAR TECHNOLOGY. B. Hoeneisen y Carver A. Mead en *Solid-State Electronics.* vol. 15, n.º 8, págs. 891-897; agosto, 1972.

HOW BIG SHOULD A PRINTED CIRCUIT BOARD BE? Ivan E. Sutherland y Donald Oestreicher en *IEEE Transactions on Computers,* vol. C-22, n.º 5, págs. 537-542; mayo, 1973.

MICROELECTRONICA Y ORDENADOR DE USO PERSONAL

TOWARDS A THEORY OF INSTRUCTION. Jerome S. Bruner. Belknap Press of Harvard University Press, 1966.

ARTIFICIAL INTELLIGENCE. Seymour A. Papert y Marvin Minsky. Condon Lectures, Oregon State System of Higher Education, 1974.

PERSONAL DYNAMIC MEDIA. Alan C. Kay y Adele Goldberg en *Computer,* vol. 10, n.º 3, págs. 31-41; marzo, 1977.

JUEGOS MATEMATICOS

CONIC SECTIONS. Henry Martin Taylor en *Encyclopaedia Britannica: Vol. 6, Ninth Edition.* Encyclopaedia Britannica. Inc., 1890.

GEOMETRY AND THE IMAGINATION. David Hilbert y S. Cohn-Vossen, traducido por P. Nemenyi. Chelsea Publishing Company, 1952.

A BOOK OF CURVES. E. H. Lockwood. Cambridge University Press, 1961.

TALLER Y LABORATORIO

COOL? E. B. Mpemba y D. G. Osborne en *Physics Education,* vol. 4, n.º 3, págs. 172-175; mayo, 1969.

THE FREEZING OF HOT AND COLD WATER. G. S. Kell en *American Journal of Physics,* vol. 37, n.º 5, págs. 564-565; mayo, 1969.

COOLER? Ian Forth en *Physics Education,* vol. 6, n.º 1, págs. 32-41; enero, 1971.

COOLER-LOWER DOWN. Eric Deeson en *Physics Education,* vol. 6, n.º 1, págs. 42-44; enero, 1971.